Bell Telephone Laboratories

AFCRL-67-0068

EFFECTS OF RADIATION ON SEMICONDUCTOR MATERIALS AND DEVICES

by

D. K. Wilson

J. P. Mitchell

J. D. Cuthbert

R. R. Blair

Bell Telephone Laboratories, Incorporated

on behalf of Western Electric Company, Incorporated

222 Broadway, New York, New York 10038

Contract No. AF19(628)-4157

Project No. 4603 Task No. 460801

FINAL REPORT

Period covered: 1 October 1964 thru 30 November 1966

31 December 1966

Bell Telephone Laboratories
Research and Development Unit of the Bell System

AFCRL-67-0068

EFFECTS OF RADIATION ON SEMICONDUCTOR MATERIALS AND DEVICES

by
D. K. Wilson
J. P. Mitchell
J. D. Cuthbert
R. R. Blair

Bell Telephone Laboratories, Incorporated
on behalf of Western Electric Company, Incorporated
222 Broadway, New York, New York 10038

Contract No. AF19(628)-4157
Project No. 4608 Task No. 460801

FINAL REPORT

Period covered: 1 October 1964 thru 30 November 1966
31 December 1966

Prepared for Air Force Cambridge Research Laboratories
Office of Aerospace Research, United States Air Force
Bedford, Massachusetts 01730

ABSTRACT

Results of investigations on the effects of nuclear radiation on semiconductor materials, device surfaces, and devices are discussed.

Radiation damage in gallium phosphide was studied using electro- and cathodoluminescence. Studies were also made of radiative and non-radiative recombination mechanisms in various compound and elemental semiconductors. A non-radiative Auger-type mechanism observed at neutral defect centers appears to explain non-radiative lifetime degradation from both chemical and radiation damage defects.

The Fermi level dependence of the ESR spectrum associated with the phosphorus-vacancy complex was studied in electron-bombarded, phosphorus-doped LOPEX silicon. These studies confirm one assumption that the Si-G8 (E) center is not seen until the Fermi level falls below E_c-0.48 eV.

Experiments to determine the effects of device bias, temperature, and radiation dose rate on surface damage to MOS FET's showed qualitative agreement with a model of positive space charge buildup at traps in the devices' SiO_2 layer. Preliminary thermoluminescence studies to determine the source of traps are described.

Transient recovery phenomena in silicon devices after bursts of fast neutrons were studied. A comparison with the cluster model for neutron damage implies that transient recovery is an electronic, not an atomic process. High-injection-level effects in irradiated transistors and PNPN devices are analyzed and compared with experimental results. Narrow-base PNPN devices are shown to be markedly superior to transistors as power switches in a radiation environment.

CONTENTS

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

ILLUSTRATIONS

ILLUSTRATIONS (continued)

ILLUSTRATIONS (continued)

ILLUSTRATIONS (continued)

ILLUSTRATIONS (continued)

ATTACHMENT II

THE EFFECTS OF FAST-NEUTRON IRRADIATION ON THE SATURATION VOLTAGE OF SILICON PLANAR EPITAXIAL TRANSISTORS

ILLUSTRATIONS (continued)

ILLUSTRATIONS (continued)

TABLES

TABLES (continued)

EFFECTS OF RADIATION ON SEMICONDUCTOR MATERIALS AND DEVICES

1. INTRODUCTION

Studies under this contract were carried out in three general areas:

1. An investigation into the nature of radiation damage defects in elemental and compound semiconductors using electron spin resonance (ESR) and optical spectroscopy.
2. A summary of previous studies of radiation-induced surface effects in semiconductor devices and a subsequent investigation into the nature of surface charge buildup in irradiated planar silicon devices.
3. An investigation into the effects of fast neutron irradiation on state-of-the-art silicon devices.

In fundamental studies of radiation-induced defects in semiconductors, many avenues of approach are possible through the variety of physical tools which one may select. The type of information that may be obtained depends to a considerable extent on the initial selection. Two broad classifications may be distinguished:

1. Techniques employing bulk electrical measurements, involving lifetime, resistivity, Hall and thermal measurements. These give valuable information on macroscopic parameters such as defect concentrations, minority carrier lifetime, and Fermi level location.
2. Microscopic techniques, such as electron spin resonance (ESR), photoconductivity, and optical fluorescence. These tend to give more detailed atomic information.

Bulk and atomic effects are closely interconnected; for instance, the Fermi level plays a crucial role in the interpretation of ESR and optical data.

Section I of this report contains a description of how ESR and optical techniques were used in the fundamental study of defects in silicon and gallium phosphide (GaP). Where necessary, electrical measurements were used as a supplement to these measurements. Although there is some degree of overlap in the information which these two tools provide about the defect state, they are mostly complementary. ESR provides information about the structure of the actual

defect level and optical techniques elucidate recombination mechanisms. While ESR is a proven and established tool in the investigation of radiation-induced effects in semiconductors, optical techniques have not been widely employed.

In the optical studies, electro- and cathodo-luminescent techniques have been used to study the electronic effects of radiation-induced defects and radiative recombination mechanisms in several different semiconductors with principal emphasis on GaP.

In the ESR studies, the production rate and dependence on Fermi level of the Si-G8 (E) center in phosphorus-doped LOPEX silicon has been carefully examined in an attempt to clarify the structure, charge state, energy level, and capture cross sections of this center.

Section II of this report deals with the investigation of surface effects in semiconductor devices resulting from exposure to radiation. The first part of the effort was spent preparing a summary of past work done in this area. The summary included a discussion of the effects of radiation on both nonpassivated Si and Ge devices and the more recent passivated Si planar devices. The current problem of space charge buildup in SiO_2 passivation layers was treated in some detail.

After the completion of this summary, an experimental program was evolved to investigate the effects of bias, temperature, and dose rate on the space charge buildup using commercial MOS-FET's. The results of these studies are compared with an analysis of a model for the charge buildup process which assumes the trapping of radiation-produced holes.

Two other studies related to surface effects were also carried out. In one, the thermoluminescence of irradiated SiO_2 layers was investigated to determine whether more information could be obtained about the hole traps in the oxide. In the other study, an attempt was made to investigate the effect of radiation on states at a metal-semiconductor interface using Schottky barrier diodes.

Section III of this report is concerned with the effects of fast neutrons on silicon devices. The principal effect of neutrons in most devices is lifetime degradation. In some cases, the loss of majority carriers due to neutron bombardment may also be a problem.

The first part of Section III is devoted to a study of short-time recovery of lifetime degradation effects in silicon devices following a transient burst of fast neutrons. In this study the implications of the cluster model for fast neutron damage are discussed, and it is pointed out that the large number of charges associated with each cluster could explain the observed recovery effects. This possibility is compared with the currently assumed model that recovery is due to short-time annealing of neutron damage.

The second part of Section III is given over to a discussion of high-injection level effects in short-lifetime junction devices. In transistors at high injection levels, emission crowding effects drastically limit the current- and power-handling capabilities. The result of these effects on the saturation behavior of power switching transistors is discussed in detail in Attachment II. Attachment III proposes a solution to the power limitations of transistors in a radiation environment by using PNPN devices. The radiation response of narrow base PNPN structures is compared to a simple device theory. Design criteria for radiation-hardened PNPN and PIN devices are discussed.

In the last part of this section there is discussed the design and construction of a logarithmic curve tracer for radiation damage studies in silicon devices.

I. STUDIES OF RADIATION EFFECTS ON MATERIALS

The work in this general area was subdivided into two principal parts:

1. Optical studies in semiconductors, principally GaP.
2. Electron spin resonance (ESR) studies in phosphorus-doped LOPEX silicon.

The optical studies in semiconductors are reported in detail in Scientific Report No. 2.[1]* Consequently, the discussion which follows will give only a summary of that report. The ESR studies of LOPEX silicon performed under this contract were discontinued in June 1966, so the results discussed below cover only work done up to that time.

2. OPTICAL STUDIES

a. Research Objectives

It was originally decided to devote most of the effort to a study of radiation damage in gallium phosphide, GaP, a material which promises to have important device applications. The ready supply of GaP within Bell Telephone Laboratories and its efficient luminescence in the visible region made it a natural candidate for radiation damage studies. The basic quantities to be obtained in the experiments were the fluorescent spectra and lifetimes as a function of radiation damage. The former give information about the densities of centers and their energy levels and the latter information on the kinetics of the recombination processes. The way the program developed and was influenced by practical considerations are implicit in the following chronological outline of the phases of work.

1. Construction of experimental facility (October 1964 - February 1966).
2. Radiation damage in electroluminescent GaP (January - April 1965).
3. Radiation damage studies in GaP performed at BTL Murray Hill (April - September 1965).
4. Use of completed BTL Whippany facility to study general radiative recombination mechanisms (February - October 1966).

*References are listed on page 108.

The initial phase of the work involved the planning and building of the electron spin resonance (ESR), and optical spectrometers, and Van de Graaff irradiation facilities. Before the installation of the Van de Graaff, gamma- and electron-induced radiation damage was studied in GaP using the electron and gamma facilities at BTL Murray Hill. Following the completion of the facilities at BTL Whippany, all further work was performed there. The latter work was centered upon studies designed to lead to a more general understanding of the way in which carriers recombine, both radiatively and nonradiatively.

The following section contains a summary of the contents of Scientific Report No. 2 in which a detailed account is given of the actual research undertaken.[1]

b. Summary of Results

(1) Radiation Damage Fulfillment of the research objectives necessitated the building of a facility, shown in Figure 1, by which damage could be introduced into materials at very low temperatures and suitable excitation could thereafter be

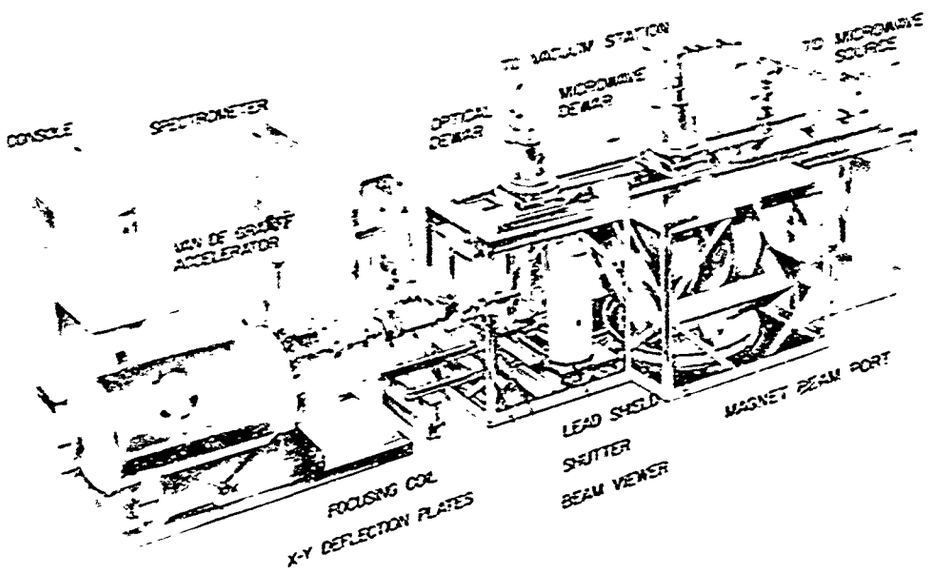


Figure 1. Van de Graaff Irradiation Facilities at BTL Whippany Laboratory

applied to the samples to study the fluorescent properties. The Van de Graaff accelerator shown in Figure 1 is capable of producing 400 kev electrons. A switching mechanism and pulser was designed, built, and incorporated into the high-voltage terminal of the accelerator. This enabled a d-c beam (0 - 100 μa) or a pulsed beam (0 - 10 ma) to be selected. Besides specialized items such as the electrostatic deflection plates, focusing coil, and electron beam shutter, special attention was given to the pumping system, so that the dewar and drift tube vacuums could be made contiguous. In this way samples could be excited while they are in the liquid refrigerant. The optical detection system consists of a grating spectrometer and photomultiplier, the output of which can be used to obtain normal or time-resolved spectra and fluorescent lifetimes. The short cutoff time of the electron pulses enabled fluorescent lifetimes as low as 4×10^{-9} sec to be measured.

Radiation damage produced at 4°K by 400 kev electrons has been studied by its effect on the Si-S pair spectra[2] in a crystal of GaP Pair spectra arise through the recombination of electrons and holes trapped at associated, ionized donor-acceptor pairs. It was found that the pair lines, which characterize the discrete separations between associated donor-acceptor pairs, are uniformly attenuated in intensity without broadening or energy shifts. The measurements were accounted for in terms of a simple theory which enabled two independent and consistent values to be obtained for η_N, the fractional production rate of nonradiative pairs. Assuming an average threshold for displacement of 10 ev, the number of vacancy-interstitial pairs in GaP per unit incident electron flux can be evaluated from the work of Oen and Holmes.[3] It is concluded that if a damage defect (presumably the phosphorous vacancy will be the most abundant) is formed within about 100 A of the "center" of a pair, then the probability that it will "kill" the fluorescence normally associated with the pair is near unity. The experimental facts are accounted for by ascribing to the defect deep acceptor characteristics and by noting that the wave functions of neutral silicon are very widespread. Therefore the holes, which are initially captured by ionized silicon, can readily tunnel to the deep acceptor defect where they are tightly bound, thus preventing the production of S-defect pair emission. The neutral defect can then act as a center for efficient nonradiative recombination through an Auger mechanism. The complete lack of any fluorescence is in keeping with the observation that in GaP neutral acceptors do not produce measurable fluorescence.

Insofar as carriers associated with radiation damage defects are nearly always tightly bound, electron hole recombination at point defects will predominate over recombination at associated impurity-defect or defect-defect type centers. This is especially true for situations where low defect densities ($\sim 10^{17}$ cc) are involved and for defect production at very low temperatures which prevent appreciable

diffusion and association in most semiconductors. However, many details of carrier recombination at point defects were, and still are, unclear, especially with respect to nonradiative decay mechanisms and to the lifetimes of such processes. In view of the general importance of these considerations in the general discussion of radiation defects in semiconductors, the authors were motivated to undertake certain experiments designed to shed light on some of the existing uncertainties. Because of the much higher level of understanding associated with chemical impurities, as compared with damage defects, the former were chosen for the study.

(2) Recombination Mechanisms. The kinetics of bound exciton recombination at isoelectronic substituents[4,5] were studied first. An isoelectronic trap[6] is formed when a host atom is displaced by an impurity from the same group of the Periodic Table. The electronegativity difference between the atoms creates an energy state within the forbidden gap. Such centers are electrically neutral and there are no weakly bound electrons prior to activation. The materials studied were GaP:N, GaP:Bi, ZnTe:O and CdS:Te.*

At temperatures where no thermal quenching occurs, the fluorescent efficiency at these centers is near 100 percent. The zero phonon line consists of a doublet which arises from j-j coupling of a $j = 3\ 2$ hole and a $j = 1/2$ electron forming two states with $J = 1$ and $J = 2$. At very low temperatures, the system is constrained to radiate from the lowest energy state corresponding to $J = 2$, so that $\Delta J = 2$ transitions are involved. Such transitions are, to a certain extent, forbidden; the degree to which they are forbidden depends upon the mixing of the $J = 1$ and $J = 2$ states. This in turn is determined by the nature of the center, whether it is an electron or hole trap.

Measurements of the fluorescent lifetimes as a function of temperature enabled predictions as to the nature of the traps to be verified and the oscillator strengths and density of traps to be calculated. The latter values compared well with independent estimates of the density of traps. The thermal activation energies of both carriers were found to be quite similar in insulating crystals of GaP:N and GaP:Bi, and they exhibited a simple quenching behavior. A more complicated situation arises in the crystals of CdS:Te and ZnTe:O (n and p type respectively) due to the incorporation of stochiometric defects. By chance it happens that Te in CdS is a deep-hole trap and O in ZnTe is a deep-electron trap, so that the more weakly bound majority carrier is ionized first in both cases, leaving behind a charged center. A dynamic exchange of free and bound majority carriers occurs and

*The work on CdS:Te was not funded under the present contract and will be reported separately.[7]

has the effect of greatly lengthening the fluorescent decay times without impairing the efficiency. At higher temperatures, the fluorescence quenches with an activation energy corresponding to the loss of minority carriers.

In contrast to the above, exciton decay at neutral donors in silicon and GaP[8] was found to be very fast, and the principle of detailed balance could not relate the measured concentration of centers and absorption cross section to the fluorescent lifetimes. The discrepancy for arsenic-doped silicon was about 8000 and for GaP:S about 700. The measured radiative efficiencies were also near the inverse of these numbers. Of the possibilities suggested (and formally discussed) to resolve the difficulties, an Auger mechanism is the only one which entirely satisfies the evidence. In this nonradiative mechanism, the exciton recombination energy is given to the third weakly bound electron (or hole) which is ejected far into the conduction (or valence) band. This mechanism is similar to internal conversion and the calculation of the percentage number of internal conversion electrons can be appropriately modified to estimate the ratio of radiative to nonradiative Auger transitions. The predicted values agree well with the discrepancies between observed and calculated luminescent lifetimes and efficiencies.

The very short fluorescent lifetimes and low efficiencies observed in GaP:Te and GaP:Se[9] can also be attributed to the occurrence of the Auger mechanism. The high fluorescent efficiencies associated with ionized donor-acceptor pairs and with isoelectronic centers can be explained by the absence of this Auger mechanism, since there is no loosely bound third carrier to which the exciton energy can be transferred.

Although various Auger recombination mechanisms have been discussed in the literature* they have seldom been invoked to explain experimental data, especially in radiation damage studies. Yet it is precisely in this field that such mechanisms should play a prominent role because radiation defects energy levels generally lie deep in the forbidden gaps so that the defects will generally be neutral, favoring Auger recombination. It is felt that Auger recombination is a much more natural explanation of nonradiative decay of electron-hole pairs than the often invoked multiple phonon cascade process.[10]

*A complete bibliography is contained in Scientific Report No. 2.

3. ELECTRON SPIN RESONANCE STUDIES

a. Resonance Spectrometer

As shown in Figure 1, the electron spin resonance (ESR) spectrometer is integrated with the Van de Graaff accelerator and the optical spectrometer in such a way that low temperature (< 4.2°K) irradiations could be carried out using the ESR dewar outside the magnet. After the irradiation, the dewar can be decoupled from the accelerator drift tube and moved into its normal position within the magnet. Ultimately it is hoped that the electron beam can be brought through the magnet pole faces and irradiations performed in-situ.

The ESR spectrometer is an X-band superheterodyne system, but it differs from the conventional system in two respects:

1. The main klystron is phase-locked to the 1100th to 1200th harmonic of an 8-MHz crystal oscillator having a tuning range of 35 kHz.
2. The local oscillator, in turn, is phase-locked (slaved) to the main klystron, using a 27-MHz crystal oscillator as the i-f reference.

Thus, excellent short- and long-term frequency stability and phase coherence in microwave and i-f channels were obtained. Moreover, once the harmonic number had been determined by means of a wavemeter, the microwave frequency could be measured with an accuracy of 1 part in 10^6 by measuring the fundamental frequency of the 8-MHz crystal. Minimum detectable spins at 1.4°K have been determined to be less than 10^{11} using a powdered phosphorus-doped silicon sample. This is within a factor of 10 of the theoretical limit.

A microwave TE_{102} cavity assembly has been designed which allows samples to be introduced and removed by means of lucite rods while the cavity remains at liquid-helium temperatures. The cavity is made of gold-plated brass and has 0.010-inch wide slits on the side and bottom. These slits serve as both optical windows and ports for electron beam irradiations. Stainless steel (0.0005-inch) windows are soldered into the Dewar walls for electron-beam irradiations.

Samples used in this cavity are typically 1.2-inches long, 0.4-inch wide and 0.030-inch thick. The sample is mounted on a copper block which minimizes microwave leakage and serves as a heat sink. Resistivity and Hall coefficient are measured for each sample by the Van der Pauw method using four Au-Sb leads welded to the edges of the sample.

b. Objectives

To put the initial objective of the ESR work into proper perspective, a review of the state of affairs when the program was undertaken is in order:

1. In phosphorus-doped float zone silicon the dominant ESR center designated as Si-G8 (E) center was attributed to the phosphorus-vacancy pair.[11,12,13]

2. Electrical measurements on the same material indicated the presence of a trap level or levels somewhere between 0.38 and 0.47 eV below the conduction band.[14,15]

3 The concentration of the ESR center and the traps increased with the phosphorus concentration, and in some isochronal annealing experiments both disappeared at the same temperature of 150°C. Thus the ESR center was believed to be associated with at least one of the trap levels near E_c-0.4 eV.[14]

4. However, discrepancies in other annealing experiments indicated that not all of the E_c-0.4 eV levels were directly related to the Si-G8 (E) center.[16]

The original task was, then, to investigate the kinetics of the formation of phosphorus-vacancy pairs, the Fermi-level dependence of their formation, and the excess carrier capture cross-section of these deep trap levels. The measurements were extended to the phosphorus concentrations in the range of 10^{16} to 10^{17} cm^3, and the electrical and ESR data were carefully correlated.

c. Experimental Results

Using the above facility, ESR and electrical measurements have been made on a variety of phosphorus-doped LOPEX silicon crystals with impurity concentrations in the range 10^{15} - $10^{17}/cm^3$. Electron irradiations with a 1 MeV Van de Graaff (BTL, Murray Hill) and gamma irradiations using a Co^{60} Gammacell were performed at room temperature.

In many of the samples (irradiated to a near-intrinsic level) a single isotropic center was observed with a g = 2.0045 ± 0.0005 with a peak-to-peak linewidth of 5.5 gauss. This line is observed at 300, 77 and 4.2°K. No hyperfine splitting was observed even at 4.2°K. None of the more than 20 centers in electron-irradiated silicon reported so far corresponds to this defect. The only comparable spectrum is one reported by Chen and Kikuchi[17] in neutron irradiated silicon and was attributed by the authors to interstitial silicon. Further information on this center, which appears to be characteristic of LOPEX crystals, will

probably require electron-nuclear double resonance (ENDOR) or uniaxial stress techniques because of its isotropic behavior.

The presence of this line in many of the samples used in this research has made it difficult to obtain reliable g-tensors on other centers because of excessive overlapping of the spectra. A conversion of the present spectrometer to a higher frequency would possibly overcome this problem.

In spite of this problem, it has been possible to study the growth of the Si-G8 (E) center under electron bombardment.

It was found that the growth of the EPR signal of the Si-G8 (E) center at 20°K is very low until the total integrated flux in electrons/cm^2 is 10 to 30 times the phosphorus donor concentration per cm^3, as reported by Watkins, et al.[11] For a 0.5 Ω-cm sample (with 1.1 x 10^{16} phosphorus/cm^3) the Si-G8 (E) center was seen after 3 x 10^{17} electrons/cm^2. For a 0.08 Ω-cm sample (with 1.2 x 10^{17} phosphorus/cm^3), the Si-G8 (E) spectrum was observed after irradiation to ~ 10^{18} electrons/cm^3. These observations and the accompanying electrical data seem to substantiate Watkins' original assumption that these defects are being formed throughout the irradiation, and the emergence of the spectrum is associated with a change in the charge state as the Fermi level moves down in the gap. In the 0.5 Ω-cm sample, the locking of the Fermi level to 0.48 eV was observed to take place at about 1.3 x 10^{17} electrons/cm^3. To explain the slow growth of the Si-G8 (E) center it must also be concluded that the Si-G8 (E) levels are controlling the bulk Fermi level.

The rate of decrease of the EPR signal of isolated phosphorus donors in the 0.5 Ω-cm sample at 4.2°K agrees with that reported by Watkins;[11] however, illumination of the sample by white light increased the donor signal by a factor of 5 to 10, indicating that substantial numbers of isolated donors were still present even after heavy irradiation. If the number of phosphorus centers removed upon irradiation in an illuminated sample are equated to the actual number of phosphorus-vacancy pairs formed (N_E), the data for both the 0.5 Ω-cm and 0.08 Ω-cm samples can be approximately fitted by the first order process,

$$\frac{dN_E}{dt} = \alpha \left(P - N_E\right) \Phi$$

where P is the phosphorus concentration, Φ is the electron flux, and α is a constant. However, α is found to be smaller in the more heavily doped sample in agreement with the electrical results reported by Stein[18] and Carter.[19]

The production rate of the Si-G8 (E) centers was observed to be lower in LOPEX than in float-zone material. This may be due to a higher production rate of very deep levels in float-zone material or possibly to the involvement of dislocations in the production of Si-G8 (E) centers.

In June 1966 the support for this study was discontinued under this contract. The study since has been continued under other support.*

d. Summary

The extremely low production rate of the Si-G8 (E) center at the early stages of bombardment appears to be adequately explained by the original assumption that the EPR center showed up only when the Fermi level was depressed below a certain level, presumably the E_c - 0.48 level observed in bulk electrical measurements. However, the result implies further that the Si-G8 (E) energy levels control the bulk Fermi level. The production rate of the Si-G8 (E) center is lower in LOPEX material than in comparable float-zone material. This may be due to a higher production rate of deep levels or possibly to the participation of dislocations in the forming of Si-G8 (E) centers. A previously unreported isotropic center at g = 2.0045 is observed in most electron-irradiated, phosphorus-doped LOPEX crystals.

Experiments are being continued to determine the Si-G8 (E) energy level and its hole and electron capture cross sections by comparing bulk electrical measurements and ESR measurements. In addition, further studies of the ESR spectra under optical excitation are being made.

*A complete report on this study will be made available to the Air Force Systems Command at a later date.

II. SURFACE STUDIES

4. SUMMARY OF SURFACE EFFECTS

Nuclear radiation is known to cause both bulk and surface damage in semiconductor devices. This damage can lead to degradation and even failure of a device. In many cases techniques are available for reducing the effects of bulk damage, with the result that surface damage becomes the more important cause of degradation due to radiation. Considerable effort has been put into radiation-induced surface effects studies in recent years, and there is now a reasonably good, although mainly qualitative, picture of the degradation processes. A summary of the work up until 1965 has been compiled and was submitted as Scientific Report No. 1.[1,2]

The most important surface effect of ionizing radiation appears to be the accumulation of charge on the device surface. This charge affects the underlying semiconductor surface and causes degradation of the device characteristics, particularly the reverse-bias leakage current of p-n junctions and the gain of transistors.

The surface charge may collect in two distinct ways. In one process, radiation produces ions in the device ambient, e.g., the gas in the transistor can, and these ions subsequently deposit charge on the device surface. For nonpassivated surfaces this charge is very close to the semiconductor and will have a profound effect on it. Nonpassivated devices are, therefore, very sensitive to a surface charge buildup of this type. Passivated surfaces, on the other hand, are not nearly as sensitive since the surface charge in this case is separated from the semiconductor by the passivation layer (usually a few thousand Å units of SiO_2). However, charge can accumulate in another way: It is now apparent that ionizing radiation affects the oxide passivation layer itself and causes a space charge to accumulate in the SiO_2 near the SiO_2-Si interface. This space charge is quite effective in producing device degradation. The effects of radiation on SiO_2 are complicated and the details of the charge buildup are not well understood at present.

The purpose of the work described in the following sections was primarily to investigate the radiation-induced charge buildup in SiO_2 passivation layers. Based on the results of these investigations, a quantitative model has been proposed which satisfactorily explains the main features of the charge buildup process. Some preliminary results of an investigation of the effects of radiation on Schottky barrier diodes are also given. This study was undertaken to investigate radiation effects on metal-semiconductor contacts.

5. THE USE OF MOS-FET's TO STUDY RADIATION-INDUCED CHARGE BUILDUP IN SiO_2 LAYERS

The properties of the SiO_2 layer used in Si planar devices have been studied extensively in recent years, with special attention to the instability shown by this layer when subjected to temperature-bias stress. It has now been established that the instability is due to ionic motion in the oxide (notably Na^+). At elevated temperatures and with a field applied across the oxide, the ions migrate to the $Si-SiO_2$ interface giving rise to a positive space charge in this region. The metal-oxide-semiconductor (MOS) structure [see Figure 2(a)]has proved to be a convenient device in which to observe the space charge buildup. The theory and uses of the MOS structure are discussed in detail in Attachment I.

The capacitance, C, of the MOS structure is voltage dependent. Curve 1 in Figure 2(c) illustrates the dependence of C on the applied voltage, V_G, at a high frequency for p-type Si (for n-type Si the curve is reversed in the voltage direction). If a positive space charge, Q_{SS}, builds up in the oxide [see Figure 2(a)], a negative charge, Q_{Si}, will be induced in the Si; i.e., the conduction and valence bands of the Si will be depressed at the surface of the Si. To restore the Si surface to its original condition it is necessary to apply a potential ΔV to the metal gate where

$$\Delta V = \frac{Q_{Si}}{C_{ox}}$$

C_{ox} is the capacity of the oxide. If we assume Q_{ss} is distributed uniformly within a distance d of the SiO_2-Si interface [see Figure 2(b)], then

$$Q_{Si} = -\left(1 - \frac{d}{2x_o}\right) Q_{ss} \text{ and } \Delta V = -\left(1 - \frac{d}{2x_o}\right) \frac{Q_{ss}}{C_{ox}}$$

x_o is the thickness of the oxide.

It is apparent then that the effect of a positive space charge, Q_{SS}, is to shift the $C-V_G$ curve of a MOS capacitor, parallel to the voltage axis, toward more negative voltages, as shown in curve 2 of Figure 2(c). Furthermore, the voltage shift is proportional to the space charge density, Q_{SS}, and hence ΔV is a direct measure of Q_{SS}.

When SiO_2 layers are exposed to ionizing radiation, Q_{SS} is observed to increase as evidenced by a negative shift of the $C-V_G$ curves. In addition, there is often a distortion of the $C-V_G$ curves, and a hysteresis appears when the curves

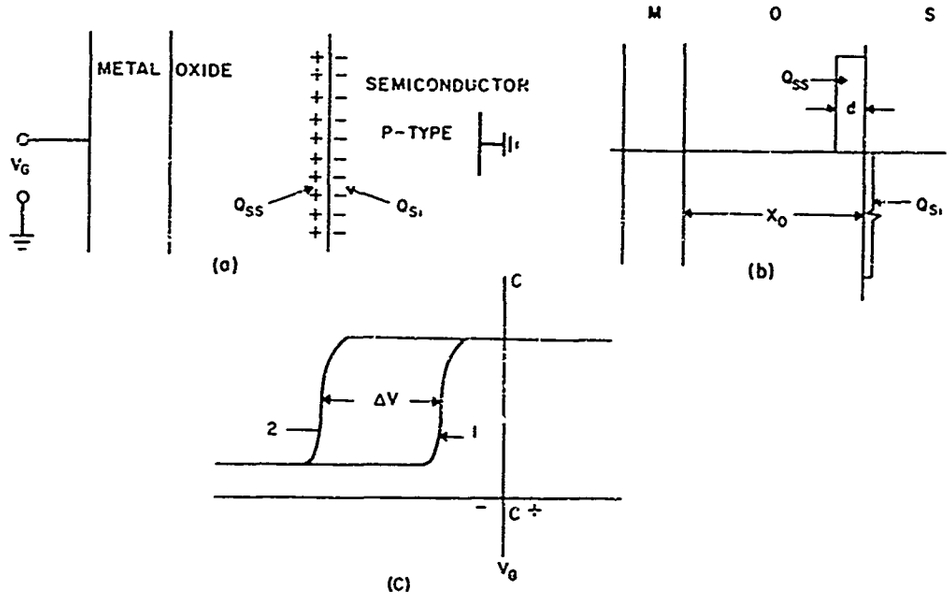


Figure 2. Space Charge Buildup in Metal-Oxide-Semiconductor Capacitor

are swept first from positive to negative and then from negative to positive V_G values. Distortion is usually interpreted in terms of electronic states at the Si-SiO_2 interface which change their charge states as the Fermi level at the Si surface passes through them. Hysteresis, on the other hand, results from a change in Q_{ss} caused by a large change of V_G from one polarity to the other.

The MOS structure is found in many bipolar transistors, usually as a result of overlaying contacts, and also in the MOS field effect transistor (FET). Commercial MOS-FET's are convenient for studying radiation-induced buildup of Q_{SS}. These devices are available with a variety of oxide types and with both n- and p-type substrates. The charge buildup can be monitored using the gate-to-substrate capacity of the device as described above, or the output characteristics can be monitored and the change interpreted in terms of Q_{SS}. The major disadvantage in using commercial units is the lack of specific information available regarding individual units.

In the experiments to be discussed in the following sections, Q_{SS} was monitored using the C-V_G curves almost exclusively. These curves were obtained manually using either a Boonton 75C variable frequency bridge (5 to 500 kHz) or a Boonton 75A-S81 bridge at a fixed frequency of 1 MHz, or automatically using a C-V_G plotter at 1 MHz as described in Appendix A.

The Raytheon RN-1030 and the Fairchild FI-100 MOS-FET's were selected as suitable devices for the study of radiation-induced Q_{SS} buildup. The RN-1030 is a p-channel enhancement mode device with a Au-Cr gate 1200A thick, 1600A of deposited oxide, and an n-type substrate. The source and drain are arranged in an interdigitated geometry. The FI-100 is also a p-channel enhancement mode device. It has an Al gate with 1200A of thermally grown oxide on an n-type substrate. Both devices have a gate-to-substrate capacitance of a few picofarads.

6. CHARGE BUILDUP

a. Factors Affecting Charge Buildup

Intuitively one might expect factors such as oxide type, temperature, radiation type, dose rate and bias applied across the oxide to have some effect on the buildup of Q_{SS}. A quantitative investigation of these factors should, therefore, yield some insight into the charge accumulation process. Accordingly, the effects of these factors were investigated.

The general procedure followed was to irradiate a device under one set of conditions, remove the charge buildup by thermal annealing, and then irradiate the same device again under new conditions. By using the same device again, it was possible to eliminate the uncertainties which result from device-to-device variations when a different device is used for each set of conditions. Of course, it is necessary to repeat the experiments several times, using different devices, to determine a typical behavior for the devices.

For the RN-1030 MOS-FET, which was used for most of the experiments, it was found that annealing at $\sim$ 300 C for 1 hour was sufficient to remove virtually all the radiation-induced space charge, even for quite badly degraded devices. Figure 3 shows two ΔV-versus-dose curves for a device irradiated with $V_G = 0$. This device was actually irradiated at a different dose rate between the two irradiations shown. Prior to each irradiation, the device was annealed as discussed above. From these results and others taken at different bias values it is apparent that the charge buildup process shows good reproducibility.

(1) Bias. The buildup of a positive space charge in a SiO_2 layer requires the movement of charge in the oxide; hence, the process by which the charge accumulates should depend noticeably on any electric fields present in the oxide during irradiation. When a bias, V_G, is applied between the metal gate and the silicon substrate of a MOS-FET, a transverse electric field appears in the oxide. The effect of V_G on ΔV was investigated for V_G in the range -15 to +10 volts.

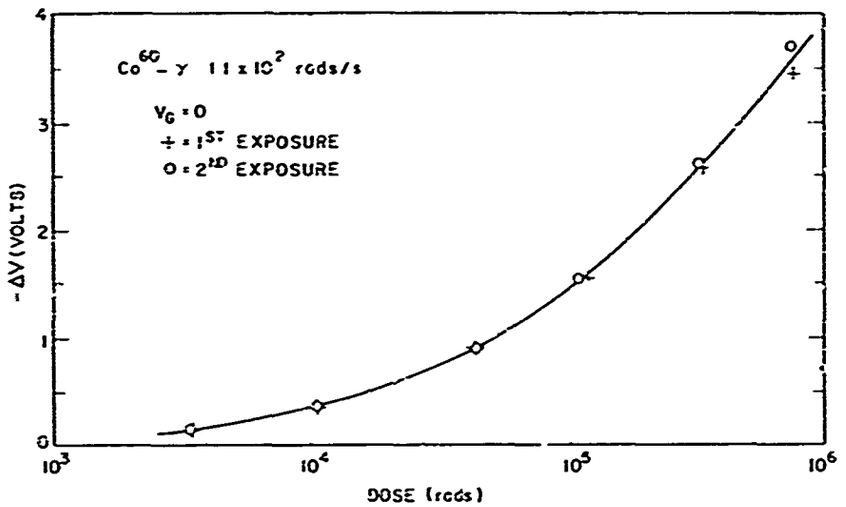


Figure 3. ΔV-versus-Dose Curves for an RN-1030 MOS-FET Irradiated Twice with $V_G = 0$

Figure 4 shows the effect of positive bias on the charge buildup process for an RN-1030 MOS-FET exposed to Co^{60}-γ radiation at a dose rate of $\sim 6.7 \times 10^3$ rads min. It should be mentioned that the doses used in this and subsequent figures are "exposed" doses, in the sense that they are the doses which would be received in the radiation chamber of the gamma cell by a properly designed dosimeter such as a LiF TLD. Because the SiO_2 layer in a MOS-FET is not a "well designed dosimeter," the actual dose absorbed by the oxide layer will be somewhat different. However, since the curves shown here are all obtained in the same manner, they can be compared without any correction for dosimetry error. This point will be discussed in more detail later.

From Figure 4 it is evident that the bias applied during irradiation has a very strong effect on ΔV. At low doses, $|\Delta V|$ increases rapidly but tends toward saturation as the dose reaches ~1 megarad. $|\Delta V|$ does not actually saturate, however, but continues to increase, even at doses as large as ~5 megarads.

Figure 5 shows results obtained on the same RN-1030 for negative values of V_G (note the change of scales). There is a definite saturation of $|\Delta V|$ for all but 0V; saturation is reached at $\sim 10^5$ rads. The two sets of curves are obviously quite different and the curve for 0V clearly belongs with the set for positive values of V_G.

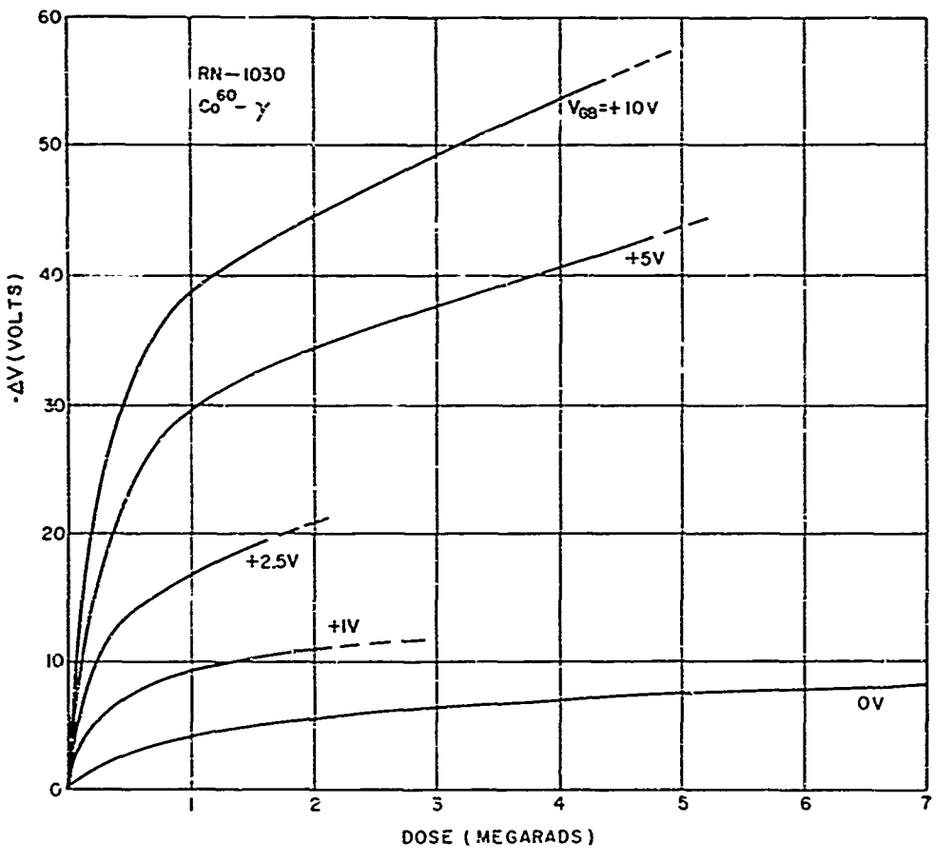


Figure 4. Effect of Positive Bias on the Space Charge Buildup in an RN-1030 MOS-FET Exposed to Co^{60}-γ Radiation

Figure 6 shows a set of curves obtained with FI-100 MOS-FET's. These curves were taken from measurements on three devices, rather than on one device as was done in the case of the RN-1030. The variation in radiation sensitivity among these devices was found to be quite small, however. Again $|\Delta V|$ is largest for positive values of V_G and does not completely saturate. For negative values of V_G the curves saturate, again at $\sim 10^5$ rads. For these devices, however, the curve for V_G = 0 volts appears to belong to the set for negative V_G values.

Figure 7 shows ΔV (sat) (defined as ΔV at 10^6 rads) as a function of V_G for both FI-100 and RN-1030 MOS-FET's. Both devices show basically the same behavior; i.e., ΔV (sat) is negative for both polarities of V_G and increases much more rapidly for positive values of V_G. There appears to be a quantitative difference between the two curves; one cause may be the difference between the methods by which the SiO_2 layers were prepared, i.e., deposited oxide in the RN-1030, and

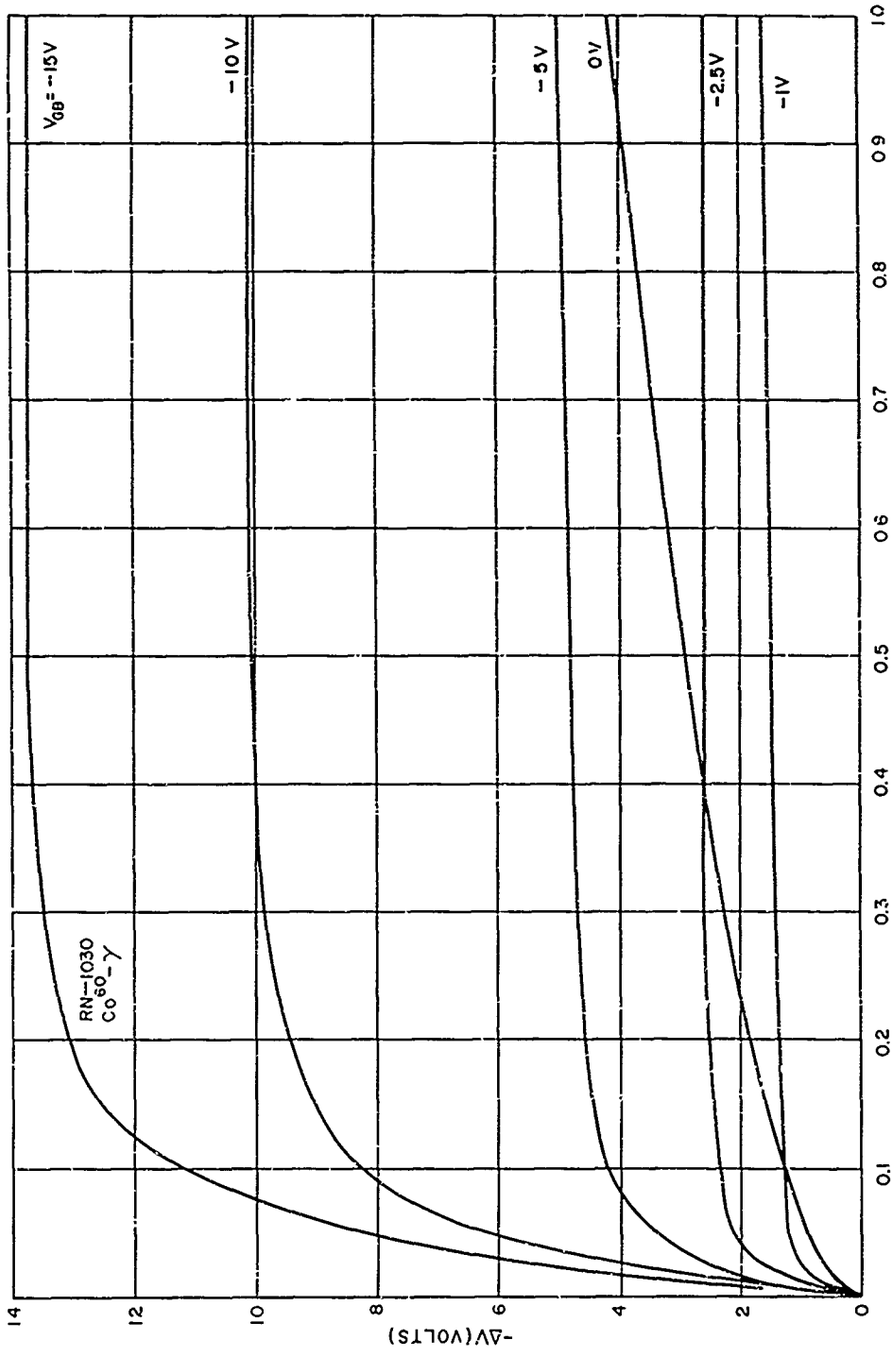


Figure 5. Effect of Negative Bias on the Space Charge Buildup in an RN-1030 MOS-FET Exposed to a Co^{60}-γ Radiation

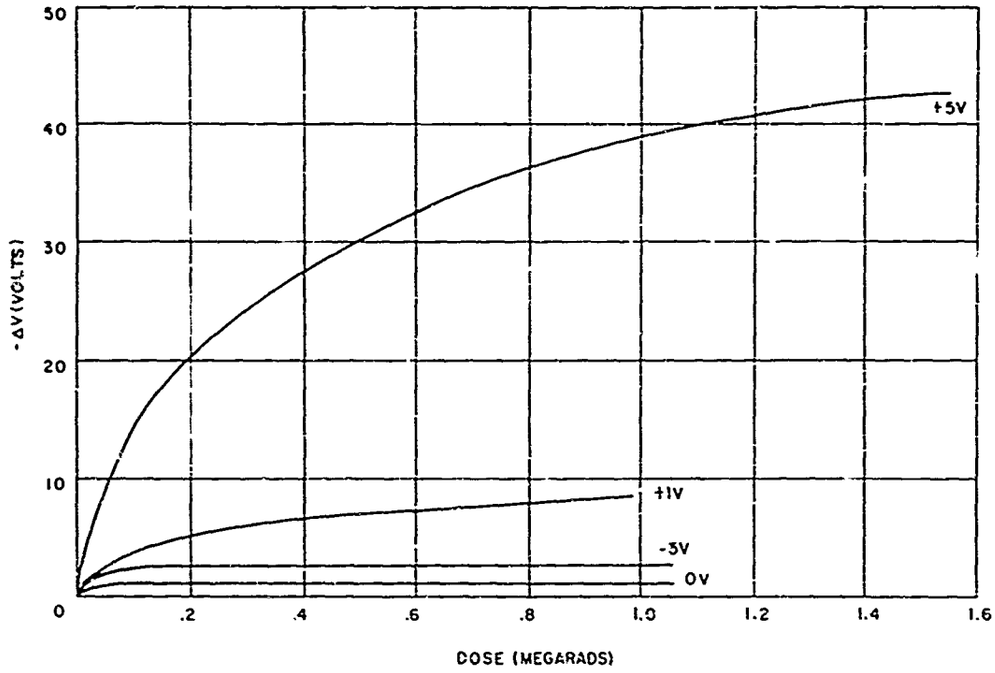


Figure 6. Effects of Bias on the Space Charge Buildup in FI-100 MOS-FET's Exposed to Co^{60}-γ Radiation

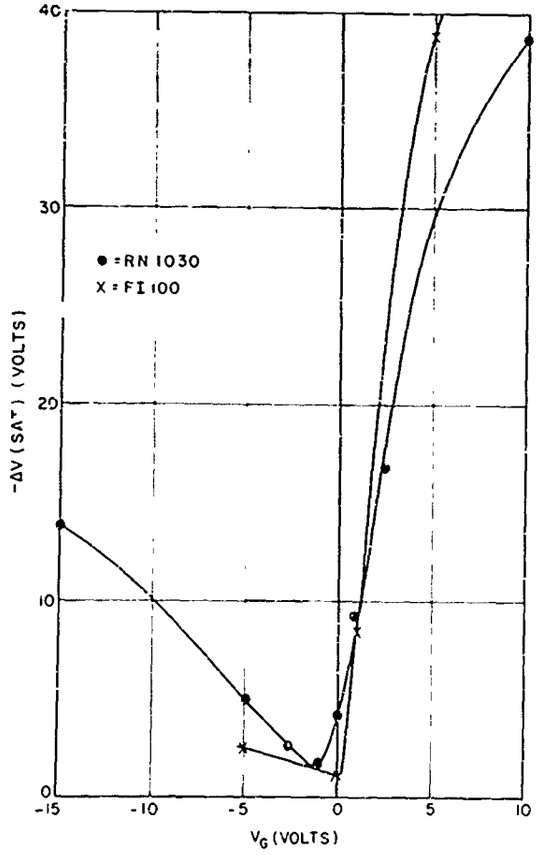


Figure 7. ΔV (sat) as a Function of V_G for FI-100 and RN-1030 MOS-FET's

thermally grown for the FI-100's. A similar difference in radiation sensitivity has been reported by Zaininger[3] for wet and dry thermally grown oxides.

Similar experiments were attempted with Motorola n- and p- channel MOS-FET's (MM 2102 and MM 2103). However, neither type of device was able to withstand more than $\sim 10^5$ rads of Co^{60}-gamma radiation before showing a high gate-to-substrate conductivity, which made it impossible to obtain $C\text{-}V_G$ curves.

(2) Dose Rate. It is not readily apparent how the charge buildup process will depend on the intensity of the ionizing radiation, i.e., on the rate at which the dose is absorbed by the oxide. This dependence is important because it may give some insight into the charge accumulation process itself. It is also important from a practical point of view, since the intensity of radiation to which semiconductor devices may be exposed can vary over many orders of magnitude, from a few rads/hr in space to 10^{10} rads/s or greater in a nuclear weapon environment. An understanding of dose rate effects is also important in simulating radiation environments and in interpreting the results of experiments carried out at different dose rates.

The dependence of space charge accumulation on dose rate was studied, using RN-1030 MOS-FET's, for dose rates in the range $\sim 10^2$ to 10^{11} rads/s. These studies were confined, for the most part, to accumulated dose levels sufficient to cause severe device degradation, i.e., $\lesssim 10^5$ rads. At these dose levels the charge buildup usually did not reach saturation.

The radiation experiments were basically of a comparative type. A device was first irradiated at the desired gate voltage using Co^{60}-gamma radiation at a dose rate of $\sim 10^2$ rads/s. The device was recovered by annealing and then irradiated again at a different dose rate, but with the same gate bias.

The dependence of charge buildup on dose rate was first investigated at dose rates of 28 and 130 rads/s using Co^{60}-gamma radiation. Figure 8 shows the voltage shift as a function of dose for bias values of 0 and 5V. Each curve is the average of results for five devices. For accumulated doses $\lesssim 10^5$ rads there does not appear to be any significant dose rate dependence for either bias value. Above 10^5 rads the curves for 0V bias diverge slightly. However, the divergence is small and may be due mainly to experimental error.

Intermediate dose rates between 10^6 and 10^7 rads/s were obtained using bremsstrahlung from a 45 MeV electron beam at the Rensselaer Polytechnic Institute Linac accelerator. The doses and dose rates were determined using LiF TLD's. The results at three bias values are shown in Figure 9. At 5V gate bias (and at 2.5 and 1V as well) the curve for the low dose rate lies above that for the high dose rate. At 0V bias the curves coincide, while at -2.5V the low dose rate curve lies below the high dose rate curve and tends to saturate at a voltage shift of $\sim$-2.5V.

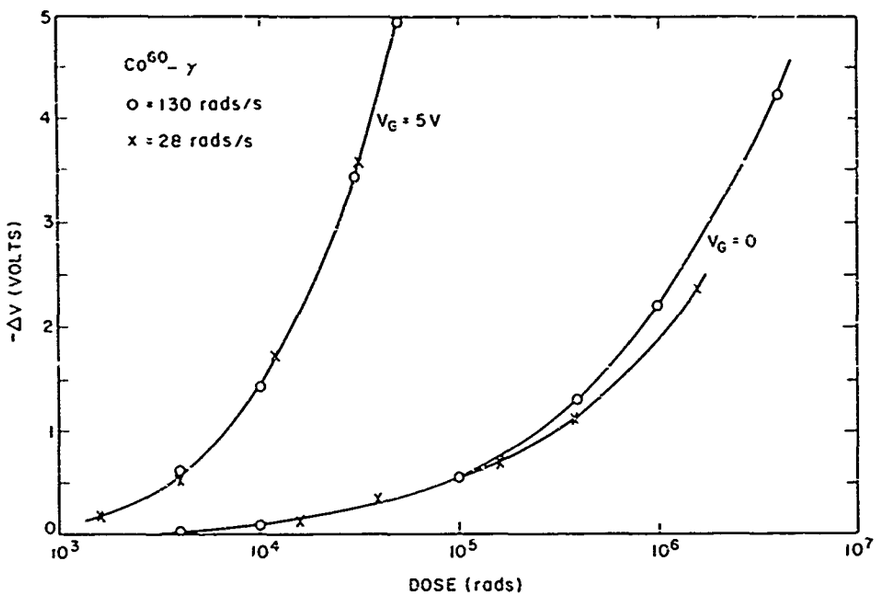


Figure 8. ΔV as a Function of Dose for Two Low Dose Rates

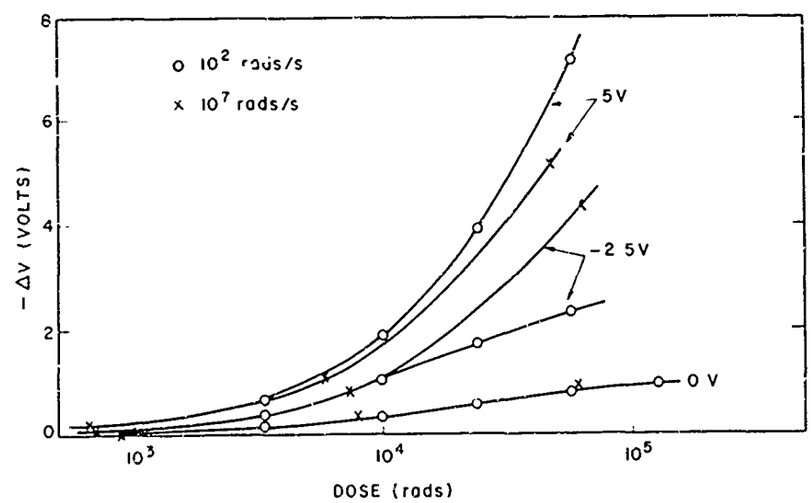


Figure 9. Charge Buildup at a Low and Intermediate Dose Rate for Three Bias Values

These intermediate dose rate results indicate that the dose rate dependence of charge buildup becomes detectable between 10^2 and 10^6 rads s. However, at least for doses $< 10^5$ rads, the effect is still quite small.

Dose rates $\sim 10^{10}$ rads s were obtained using 20 keV electrons. The electron gun used produced 2 μs pulses with current densities of $\sim$10 ma cm^2. According to the range-energy relation of Holliday and Sternglass[4], 20 keV electrons have a range of $\sim 3.7\mu$ in Si or SiO_2. Electrons of this energy will, therefore, penetrate the gate and oxide layers of a device and pass well into the Si substrate.

Figure 10 illustrates the buildup of charge with dose for devices exposed to Co^{60}-gamma radiation at $\sim 10^2$ rads s and 20 keV electrons at $\sim 10^{10}$ rads s. The dose scale for electrons in e cm^2 has been arbitrarily positioned to make the curves agree at low dose levels. With this in mind, it should be noted that at 4.6V bias the curve for the high dose rate lies below that for the low dose rate, while the opposite is true at 0V bias. Figure 11, which has the same relationship between dose scales, shows results at ±2.5V bias. At +2.5V the high dose rate curve is below the low dose rate curve, while at -2.5V the curves cross with the low dose rate curve saturating at $\sim$-2.5V. There appears to be at least a qualitative similarity between the results at this dose rate and those at the intermediate dose rate.

To compare the curves for Co^{60}-gamma and 20 keV electron irradiations properly it is necessary to obtain a conversion factor from e cm^2 to rads. On the basis of a simple model it is estimated that the oxide of the devices used absorbed $\sim 4.6 \times 10^{-8}$ rads per incident 20 keV electron. This calculation is based on depth-dose curves for 20 keV electrons and takes account of backscattering from the device surface.

As was mentioned earlier, the Co^{60}-gamma doses quoted are absorbed doses in the sense that they are the doses which would be recorded by a properly constructed dosimeter. The SiO_2 layer in a device is not well designed from a dosimetry point of view, since it is covered with higher density materials (Au and Cr) which, because of their higher conversion efficiency, will cause the SiO_2 to absorb a somewhat higher dose than would be measured by, for instance, a LiF TLD. An estimate of the size of this effect was made by placing the crystals from LiF TLD's between sheets of Au and Al of appropriate thickness to simulate a device structure, and the comparing the dose recorded by this "device" with that measured by a conventional LiF TLD. The LiF "device" recorded doses $\sim$1.4 times larger than the TLD.

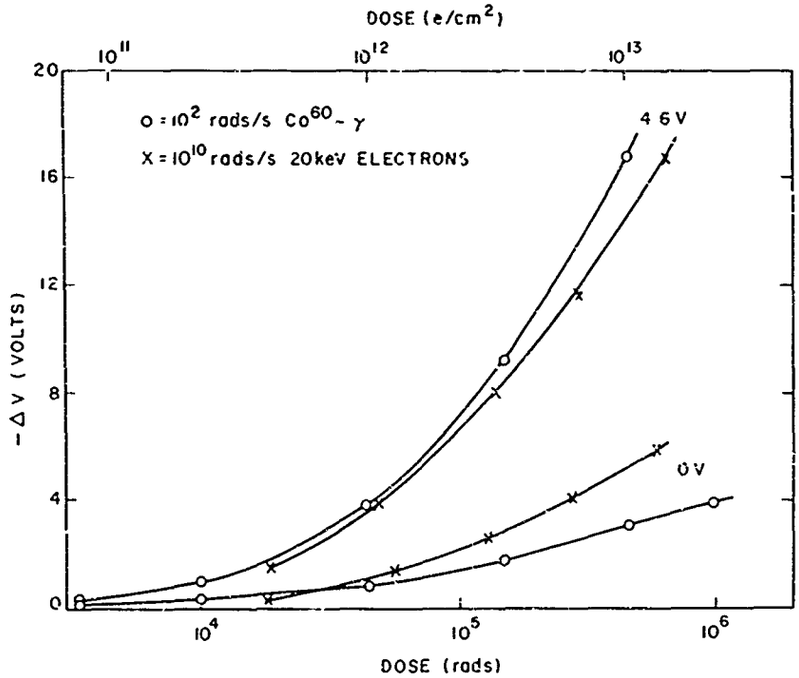


Figure 10. Charge Buildup at a High and Low Dose Rate for Bias Values of 0 and 4.6 Volts

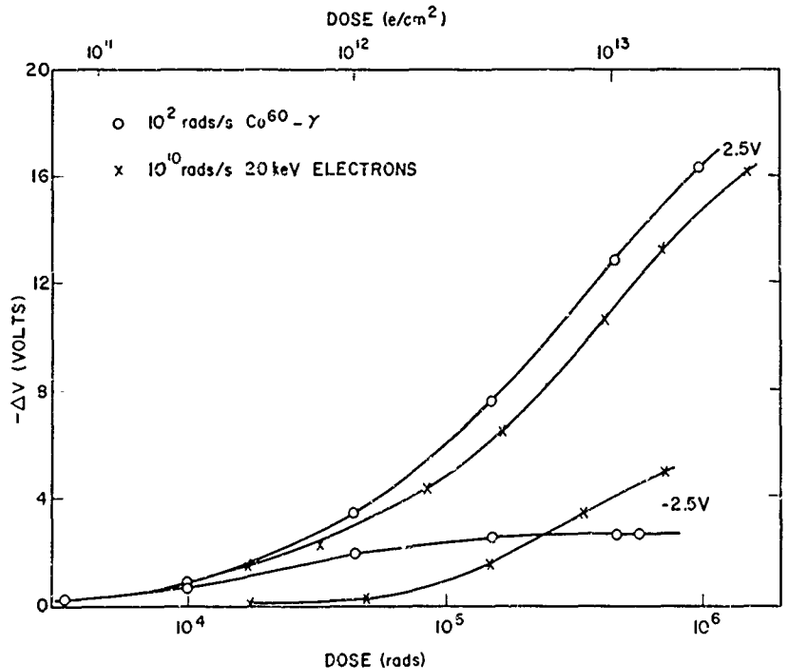


Figure 11. Charge Buildup at a High and Low Dose Rate for Bias Values of ±2.5 Volts

The results of Figure 10 are shown again in Figure 12 after the correction and conversion factors have been applied. The agreement at 4.6V (and 2.5V as well, not shown) is now quite good, while that at 0V (and -2.5V) is somewhat worse.

Dose rates approaching 10^{11} rads/s were also investigated, again using 20 keV electrons. At this dose rate it was found that the conductivity of the oxide becomes large enough so that V_G may drive appreciable current through the oxide. If V_G is a high impedance source ($\geq 10^6$ ohms), the potential across the oxide may drop significantly with the result that less space charge is accumulated.

The experiments have shown that a small dose rate dependence for charge buildup does exist in the eight decades of dose rate from 10^2 to 10^{10} rads/s. For doses less than a few times 10^5 rads, the effect is of the same order of magnitude as the variation of radiation sensitivity found in Si planar devices.

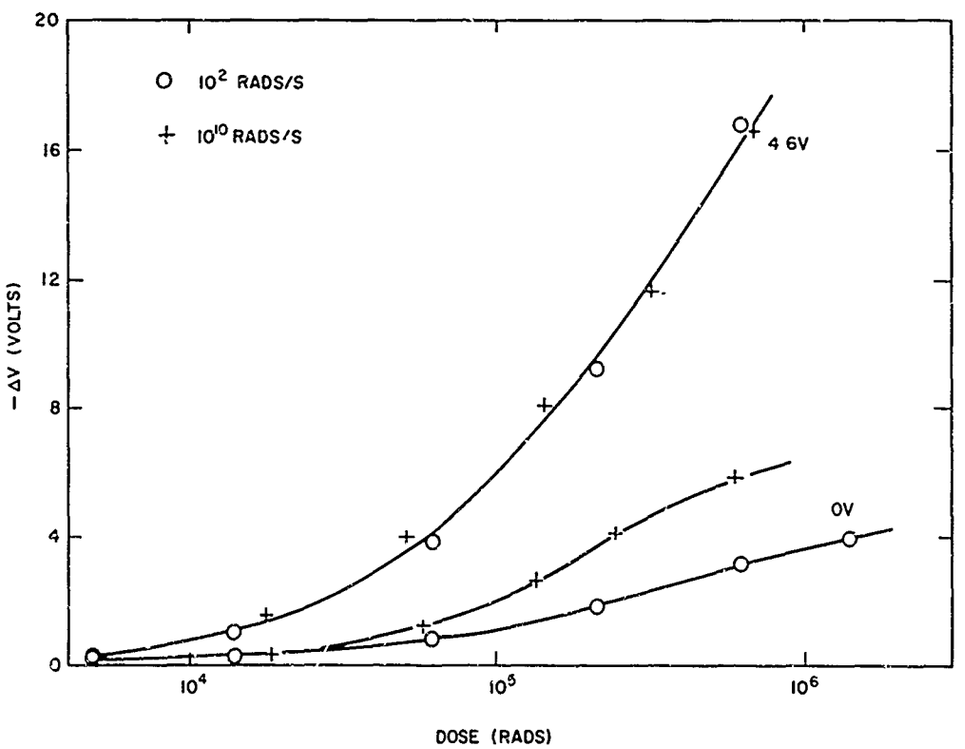


Figure 12. Charge Buildup at High and Low Dose Rates for Bias Values of 4.6V and 0V After Correction

(3) Temperature. As was mentioned previously, the radiation-induced charge, Q_{ss}, in the RN-1030 MOS-FET can be removed by a one-hour anneal at 300 C. The annealing process was studied further. Three devices were irradiated to a dose of 10^6 rads, each under a different gate bias (0, ±2.5 Volts), and then given an isochronal anneal (25C deg. steps from 50 C to a maximum of 300 C, 10 minutes at each step) with V_G=0 volts (i.e., gate shorted to substrate). C-V_G curves were taken, at room temperature, after each step. The devices were then annealed for one hour at 300 C, irradiated again under the same conditions, and then given another isochronal anneal, this time with +2.5V gate bias. The irradiation-anneal cycle was repeated again with a bias of -2.5V during the isochronal anneal.

The results for the three devices are shown in Figure 13. Since three different devices are involved, it is not possible to make quantitative comparisons. However, it is quite apparent that a large portion of Q_{ss} is removed by raising the temperature of the oxide to ~250 C regardless of the bias applied either during irradiation or during the anneal. Also, with one exception, a bias of 0 or +2.5V during irradiation is somewhat more effective in removing Q_{ss} than is a bias of -2.5V. For the device irradiated at -2.5V bias and annealed at +2.5 bias, Q_{ss} decreases until the temperature reaches ~165 C, at which point it begins to increase again. Zaininger[3] has reported similar behavior in irradiated MOS capacitors and attributed the effect to positive ion drift in the oxide. In his case, however, the positive ion instability set in at ~350 C, much higher than the temperature of onset observed here.

For all three devices the anneal with -2.5V bias gave the least recovery; in fact, a sizable fraction of Q_{ss} still remained in the devices irradiated at 0 and -2.5V bias. However, for the device irradiated at +2.5V bias, $|\Delta V|$ at 10^6 rads was initially much larger (19.1V as compared with 3.76V and 4.11V for the devices irradiated at 0 and -2.5V respectively), so that the amount of charge actually left in the oxide was approximately the same for the three devices. Annealing at -2.5V bias thus appears to hold a small residue of positive charge in the oxide. The size of the residue is approximately independent of preannealing magnitude of Q_{ss}. A one-hour anneal at 300 C and 0 bias was required to remove Q_{ss} entirely.

The results for the device irradiated at +2.5V bias and annealed at 0V bias have been replotted in Figure 14. The logarithm of the fraction of Q_{ss} annealed (i.e., log $(1-\Delta V(T)/\Delta V(297))$ is plotted against 1/T. The curve is not linear, indicating that there is no simple activation energy associated with the annealing process. However, there is a linear section whose slope yields an activation energy, ~0.15eV.

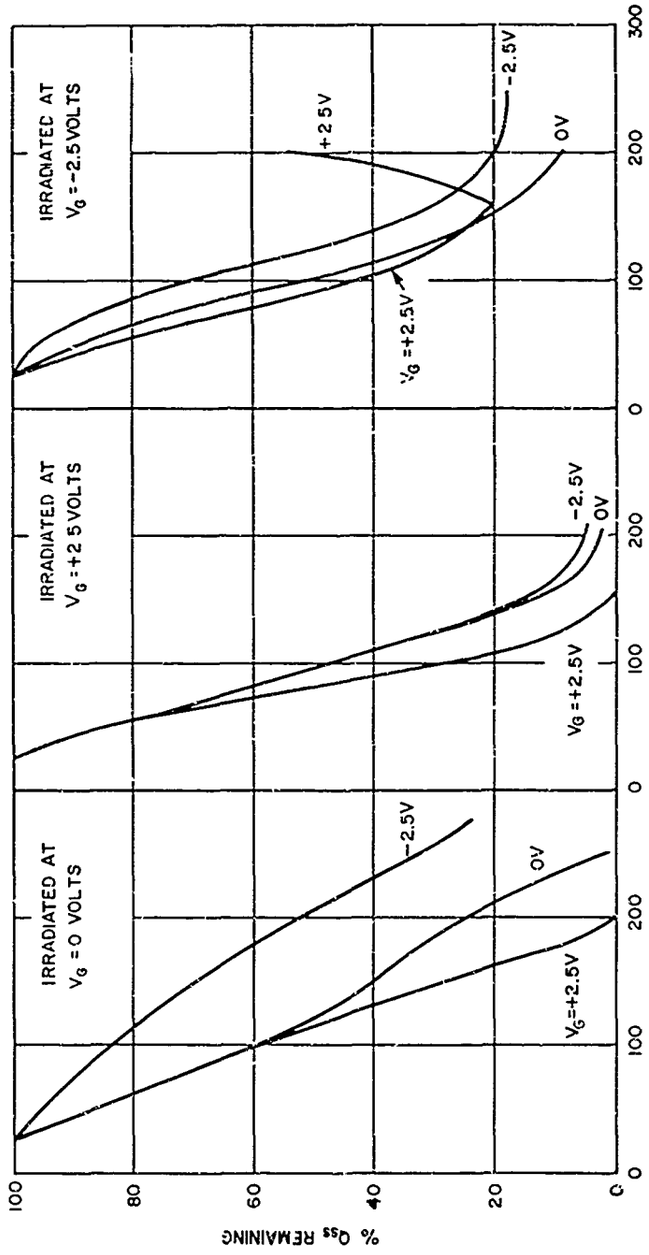


Figure 13. Results of Isochronal Annealing at Three Bias Values for Devices Irradiated at 0, ±2.5V Bias

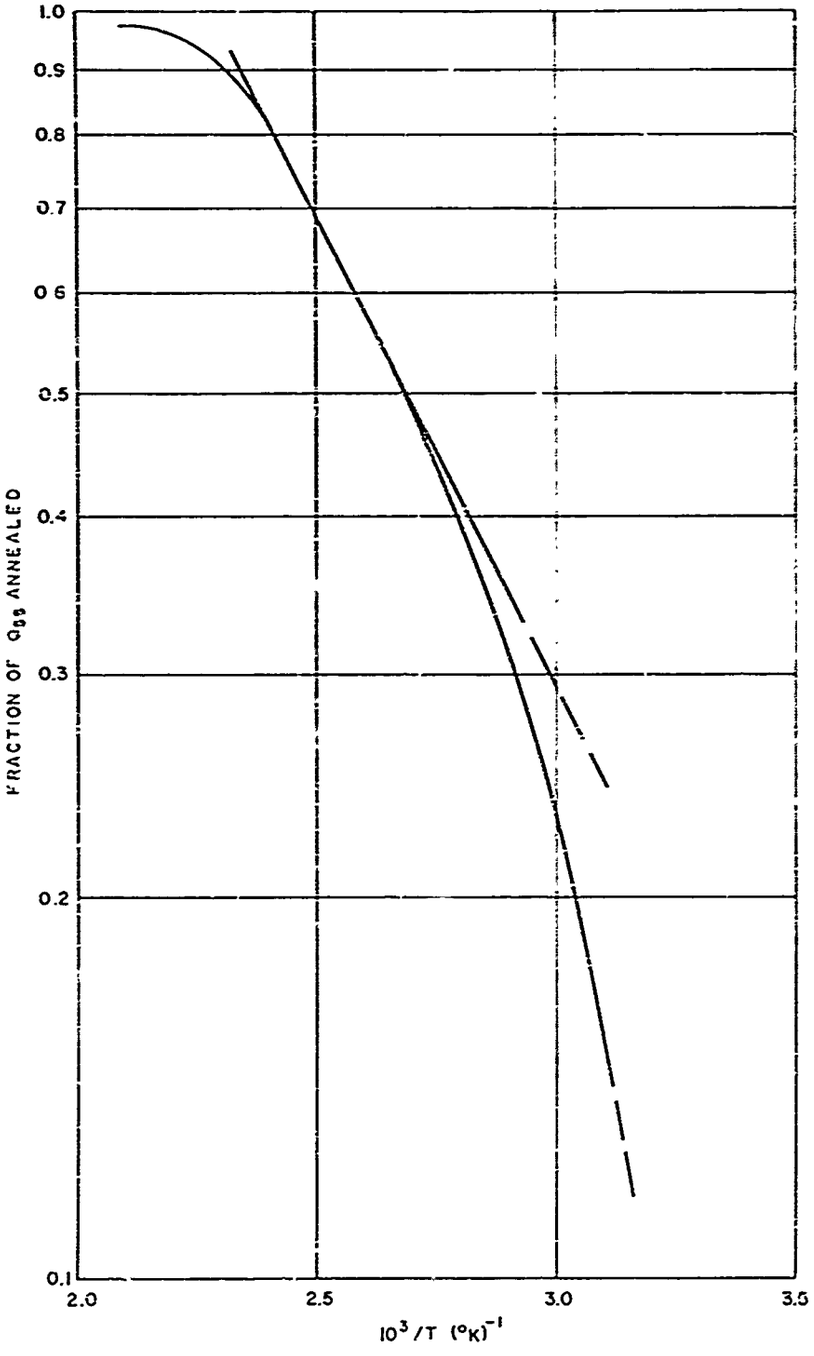


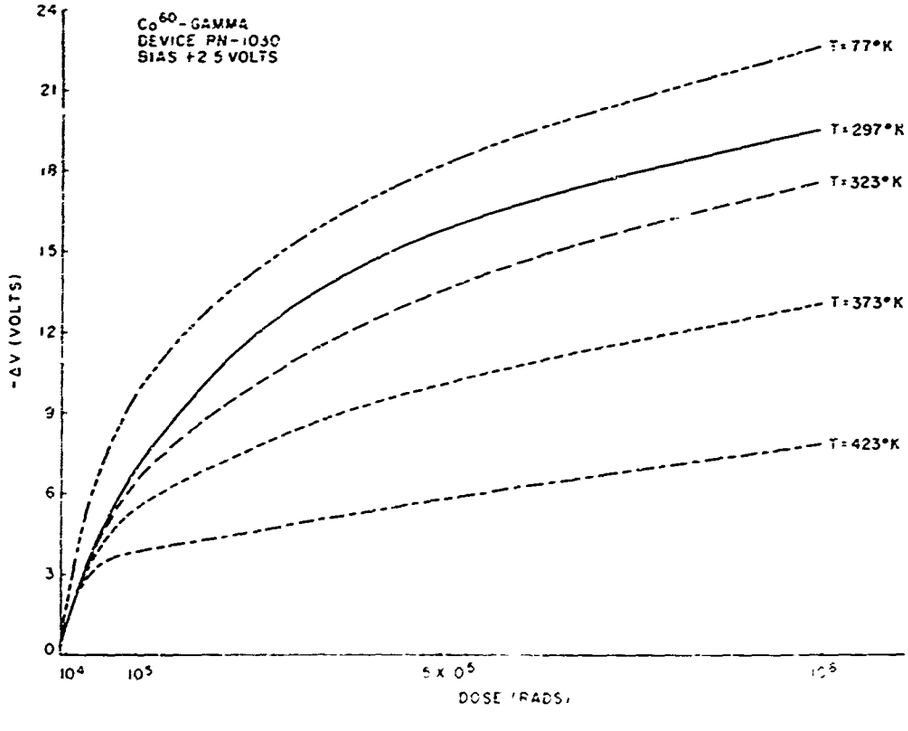
Figure 14. Isochronal Annealing Curve at 0V Bias for Device Irradiated with +2.5V Bias

Since Q_{ss} can be annealed at quite low temperatures, it is reasonable to expect that the buildup of Q_{ss} will depend rather markedly on the oxide temperature during irradiation. To determine the importance of oxide temperature, a series of irradiations was carried out on RN-1030 MOS-FET's under various bias conditions and at several temperatures. The results of these experiments may prove useful for two reasons. First, in practice, devices may be exposed to radiation at temperatures other than room temperature (the temperature at which most experiments are performed). It will be necessary, therefore, to be able to correct for temperature when making estimates of radiation effects. These experiments may indicate the importance of these corrections. Second, the results will, hopefully, yield some information about the charge buildup process itself.

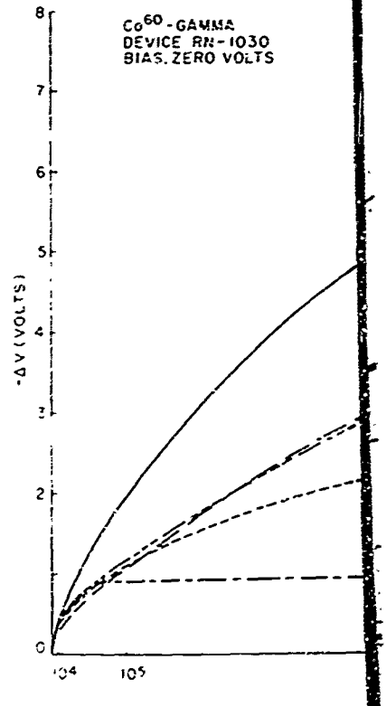
Three RN-1030 MOS-FET's were exposed at a selected temperature to a total of 10^6 rads of Co^{60}-gamma radiation. Each device was kept at a different bias (0, ±2.5V). C-V_G curves were taken, at room temperature, after accumulated doses of 10^4, $5x10^4$, 10^5, $5x10^5$ and 10^6 rads. After each curve was taken, the devices were brought to the desired temperature before the irradiation was resumed. When the accumulated dose reached 10^6 rads, the devices were annealed at 360°C for one hour and then irradiated again at a different temperature but with the same bias values. Irradiations were performed at 77, 297, 323, 373, and 423°K. For the irradiation at 77°K the devices were immersed in a liquid nitrogen bath in the irradiation chamber. For the irradiations at and above room temperature the devices were placed in a special temperature chamber mounted in the radiation chamber. The radiation doses were measured with LiF TLD's.

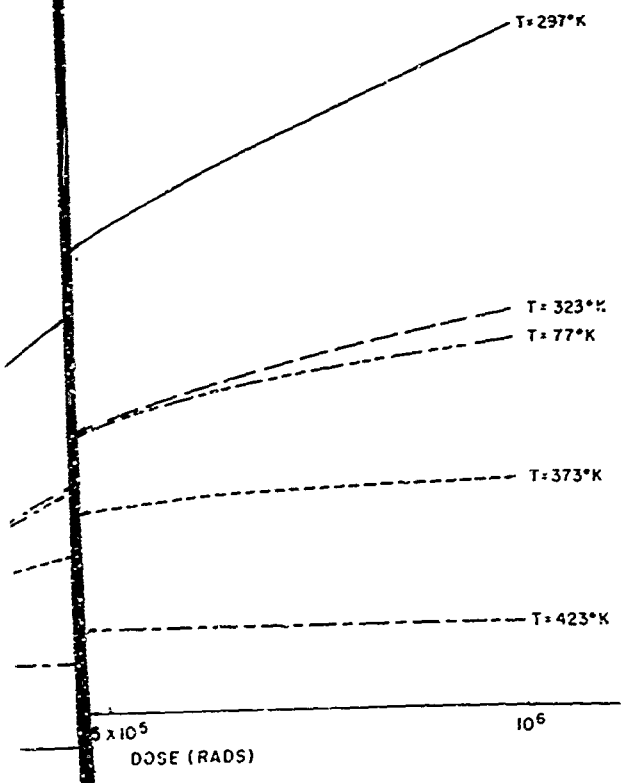
A check was made to determine the effect on the final value of ΔV of interrupting the irradiation several times and quenching the devices to room temperature. The effect should be most important at high temperatures. Three devices were irradiated, under bias, at 423°K to 10^6 rads without interruption and the results compared with those for the previous case. For each of the three cases the voltage shift, ΔV, was the same, within the limits of accuracy of the experiment, as the shift obtained when the irradiations were interrupted.

The results for the three devices are shown in Figure 15. For all three bias values, the effect of increased temperature during irradiation is to reduce Q_{ss}. The reduction as a function of temperature is illustrated in Figure 16 for a dose of 10^6 rads. Since a different device was used for each bias, it is not possible to make any quantitative comparisons. The slopes of the three curves are, however, similar to each other and also to the slope of the annealing curve in Figure 14. This fact indicates that an elevated temperature during irradiation simply superimposes an annealing process on the charge accumulation process.

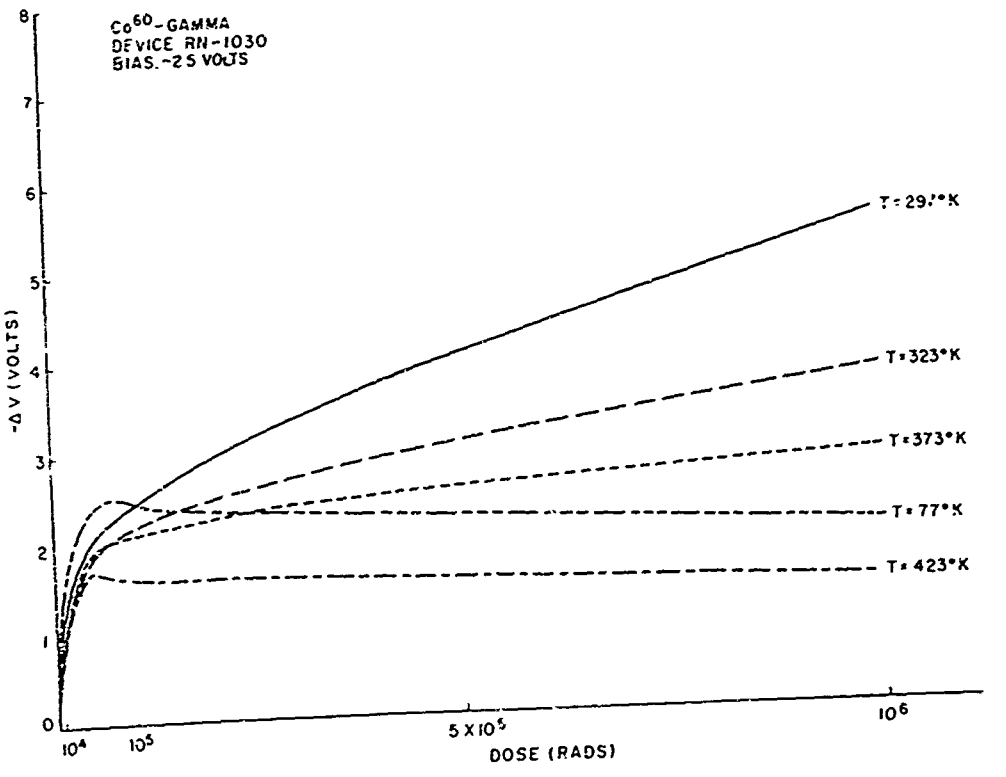

Co60 - GAMMA
DEVICE PN-1030
BIAS +2.5 VOLTS
T=77°K
T=297°K
T=323°K
T=373°K
T=423°K
-ΔV (VOLTS)
DOSE (RADS)
24
21
18
15
12
9
6
3
0
10⁴
10⁵
5 x 0⁵

(a)


Co60 - GAMMA
DEVICE RN-1030
BIAS, ZERO VOLTS
-ΔV (VOLTS)
8
7
6
5
4
3
2
0
10⁴
10⁵



(b)



(c)

Figure 15. ΔV as a Function of Dose for Several Temperatures for Devices Irradiated with 0, ±2.5 Volt Bias

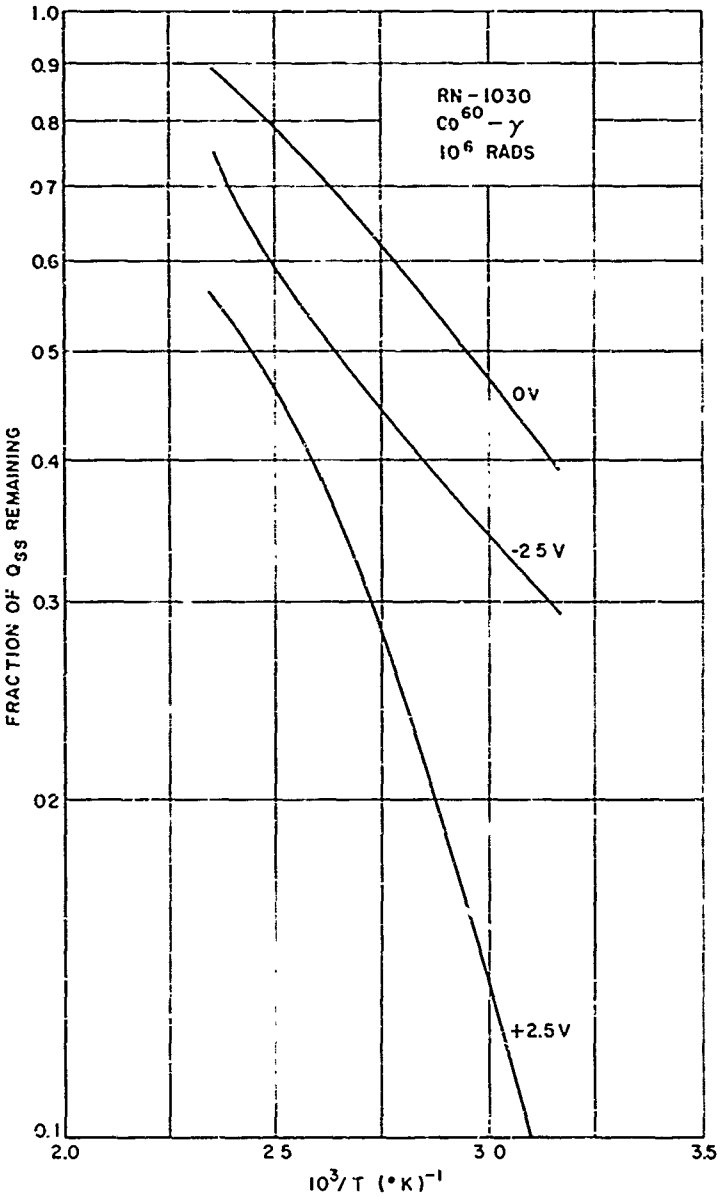


Figure 16. Annealing of Q_{ss} as a Function of Temperature for Bias Values of 0, ±2.5 Volts

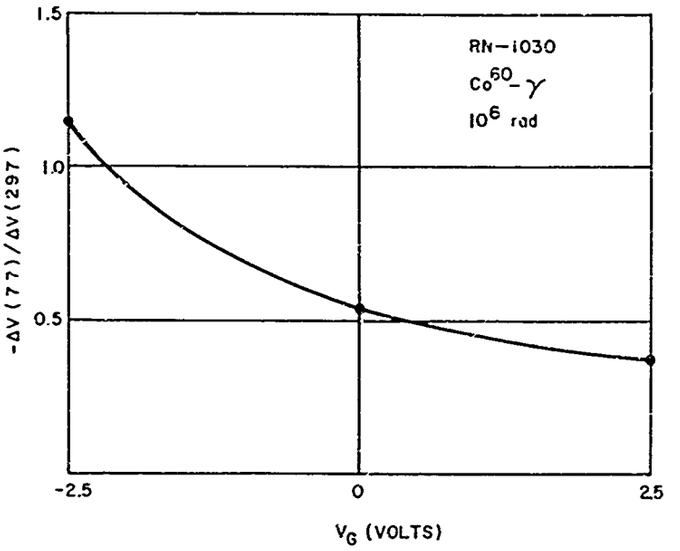


Figure 17. Bias Dependence of Charge Buildup at 10^6 rads for Devices Irradiated at 77°K

6. The charge buildup process does not depend significantly on the dose rate over a wide range of values.
7. The space charge may be removed partially or completely by thermal annealing. Annealing temperatures ~300 to 350°C are generally required for complete removal.

Isochronal annealing indicates there is no simple activation energy associated with the annealing process, but it is not unreasonable to view the process as having perhaps several such energies each a fraction of an eV. It has also been reported that the space charge can be annealed by exposing the oxide to ultraviolet light of sufficiently short wavelength.[3,6] The buildup process itself also depends on the temperature of the oxide during irradiation; the space charge accumulation decreases with increasing temperature.

Early in the investigations of the effects of ionizing radiation on SiO_2 layers, it became clear that the space charge buildup was not directly connected with ionic motion (e.g., Na^+) in the oxide. Samples which show no ion drift under thermal-bias stress do show a charge buildup when exposed to radiation.[7] Furthermore, if ionic motion were involved in the process, one would expect the effect to increase rather than decrease with elevated temperatures.

The current picture of the charge buildup process may be understood by referring to Figure 18.[3,6,7,9] Ionizing radiation passing through the oxide creates hole-electron pairs. The oxide is assumed to contain hole traps which capture some of the relatively immobile holes, while the remainder presumably recombine. The electrons, on the other hand, are more mobile and, as a result, drift toward the positive electrode. Some electrons are thus removed from the oxide and, since the Si is unable to supply electrons to the SiO_2, a net positive space charge builds up near the SiO_2 - Si interface. As the positive space charge grows, the field, E, in the oxide between the space charge and the positive electrode decreases. When the field in this region is reduced to zero, no further charge will accumulate unless the applied potential is increased. It is apparent then that the charge buildup saturates with increasing dose and, furthermore, that the saturation value will depend on applied bias, V_G.

Grove and Snow[7] and Speth and Fang[9] have proposed similar models which relate quantitatively the saturation value of the space charge to the bias applied across the oxide during irradiation, V_G. These models do not, however, explicitly discuss the charge buildup process in a quantitative way.

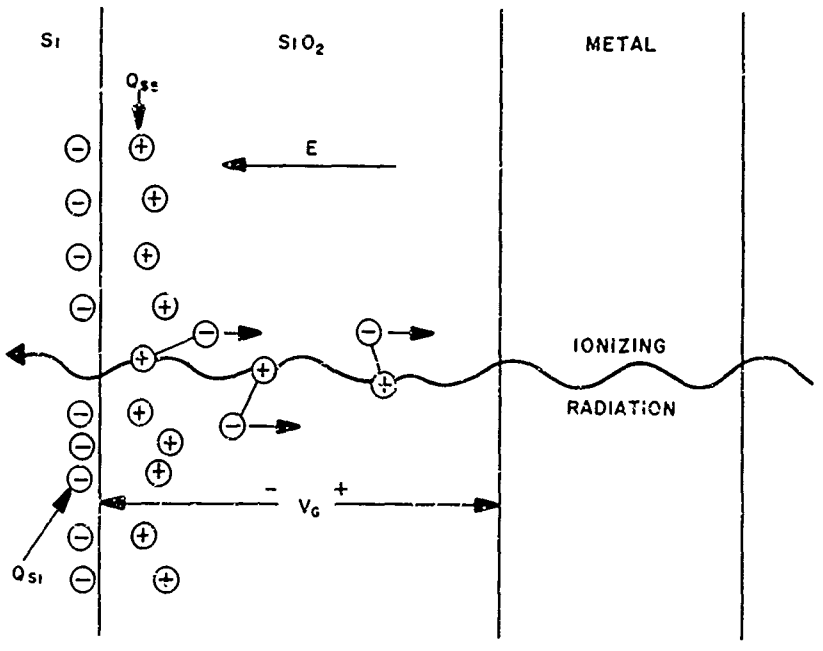


Figure 18. Model of Charge Buildup Process Due to Ionizing Radiation

(2) Quantitative Model. The picture of the charge buildup process presented in the previous section is a rather simple one and may, for this reason, yield to a straightforward analysis. In this section an attempt will be made to derive an expression for ΔV, as a function of radiation dose, which will contain the applied bias, V_G, and dose rate as explicit parameters.

It will be assumed that the oxide layer is uniform in all necessary properties and initially free of any space charge. Factors such as contact potential differences and the effect of radiation on fast states at the SiO_2 - Si interface will be ignored. Consider the situation illustrated in Figure 19. The potential, V_G, applied across the oxide gives rise to an electric field, E, which is uniform throughout the oxide and of magnitude $V_G\ x_o$ where x_o is the thickness of the oxide.

The effect of ionizing radiation is to generate hole-electron pairs uniformly throughout the oxide at a rate of g pairs cm^3-s. The effect of the electric field will be to drive electrons toward the metal gate and holes toward the Si. It is assumed that the holes are relatively much less mobile than the electrons[10] and thus a significant concentration gradient will be set up only for electrons. It is shown in Appendix B that the free electron concentration, n(x), will reach an equilibrium given by

$$n(x) = g\tau_e \left[1 - \exp\left(-a\,\frac{x}{x_o}\right)\right] \qquad (1)$$

where $a = \dfrac{x_o^2}{\mu\tau_e V_G}$, provided there has been no appreciable accumulation of space charge to alter the field, E. μ and τ_e are the mobility and lifetime of the electron respectively.

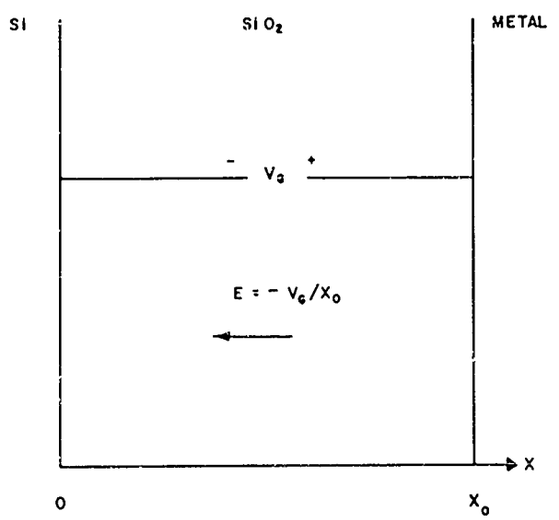


Figure 19. Initial Conditions in SiO_2 Layer

The form of the expression for n(x) is shown in Figure 20 for three values of the parameter a and for x in the range $0 \leq x \leq x_o$. The quantity a is essentially the ratio of the thickness of the oxide to the distance an electron will drift in a lifetime, τ_e. If $V_G = 0$, then a = ∞ and $n(x) = g\tau_e$ everywhere in the oxide. Experimentally it is observed that a relatively small space charge accumulates when $V_G = 0$. This charge accumulation could result either from a built-in field present in the oxide due to a contact potential, or from some other process such as diffusion. It will be assumed that in the region in which $n(x) = g\tau_e$ no net trapping of holes occurs Since n(x) is least at x = 0, the trapping rate will be greatest near the SiO_2 - Si interface. Furthermore, since it is observed experimentally that the space charge accumulates almost entirely in the small fraction of the oxide adjacent to this interface, it is concluded that a in equation 1 is ≫ 10, i.e., that $n(x) \sim g\tau_e$ throughout most of the oxide. As a result, no charge is accumulated in the bulk of the oxide and the electric field is uniform at all times except near the interface. As the space charge accumulates, the field in the bulk of the oxide decreases. When sufficient charge collects, E vanishes and no more electrons drift out of the oxide into the metal gate.

The distribution of space charge in the oxide at any time during the irradiation will depend on the density of hole traps, N_T, and the applied bias, V_G. Clearly the region in which trapping can occur, i.e., the region in which $n(x) < g\tau_e$, will increase as V_G increases, since a in equation 1 is inversely proportional to V_G. However, the maximum trapping rate occurs at x = 0 for all values of V_G. As the space charge accumulates, the field near the interface increases and maintains the high trapping rate in this region (by removing electrons). At the same time, the field in the bulk of the oxide decreases causing n(x) to approach $g\tau_e$ everywhere except near the interface. The effect will be to concentrate the trapped charge as close to the interface as possible and to prevent the space charge from spreading into the oxide.

It has been tacitly assumed that there are sufficient hole traps available close to the interface to accommodate the required space charge. If this is not the case then, of course, the space charge will spread into the oxide. The two cases are illustrated in Figure 21. N_T is the density of hole traps, assumed uniform over the region of interest, and $P_T(x)$ is the density of trapped holes as a function of distance into the oxide.

To simplify the analysis it will be assumed that at any time (dose level) the space charge distribution in the oxide can be approximated by

$$P_T = P_T(t), \quad 0 \leq x \leq d$$

$$= 0, \qquad d \leq x \leq x_o$$

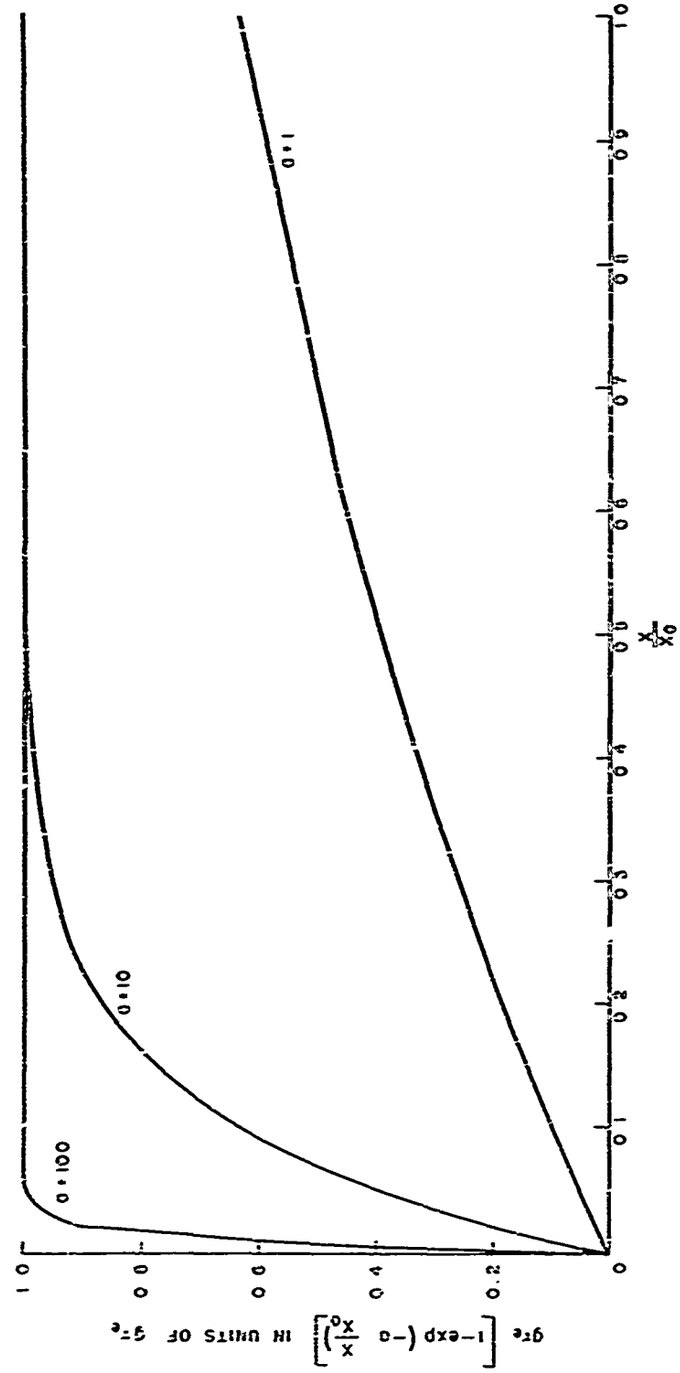


Figure 2C. $a(x)$ as a Function of x for Three Values of the Parameter a

The distribution, illustrated in Figure 21, will contain the same amount of charge as the actual distribution and have its center of gravity located at the same distance from the interface, $x = 0$.

The voltage shift of the C - V_G curve at any time, t, is given by

$$\Delta V(t) = \frac{Q_{Si}}{C_{ox}}$$

$$= -\frac{q}{C_{ox}} \int_0^d \frac{x_o - x}{x_o} P_T(t)\, dx$$

$$= -\frac{qP_T(t)}{C_{ox}x_o}\, d\left(x_o - \frac{d}{2}\right) \qquad (2)$$

It is now necessary to obtain an expression for $P_T(t)$.

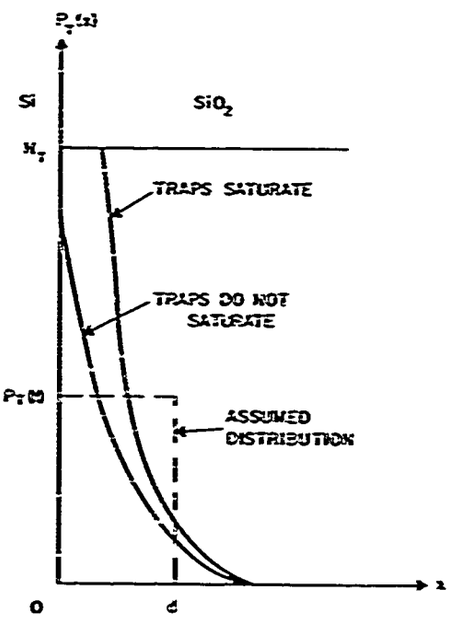


Figure 21. Actual and Assumed Distribution of Space Charge

The space charge, $Q_{SS}(t)$, is given by

$$Q_{SS}(t) = qP_T(t)d = -\int_0^t J dt \qquad (3)$$

where J is the current density due to the electrons crossing the oxide-metal interface. (See Figure 22.) If E is the electric field at the interface, then

$$J = -qg_{re}\mu E \qquad (4)$$

It is shown in Appendix C that E is given by

$$E = -\frac{V_G}{x_o} + \frac{qd^2}{2\epsilon\epsilon_o x_o} P_T(t) \qquad (5)$$

Combining equations 3, 4 and 5 and differentiating yields

$$\frac{dP_T}{V_G - aP_T} = bg\,dt \qquad (6)$$

where $a = \frac{qd^2}{2\epsilon\epsilon_o}$ and $b = \frac{\mu r_e}{x_o d}$

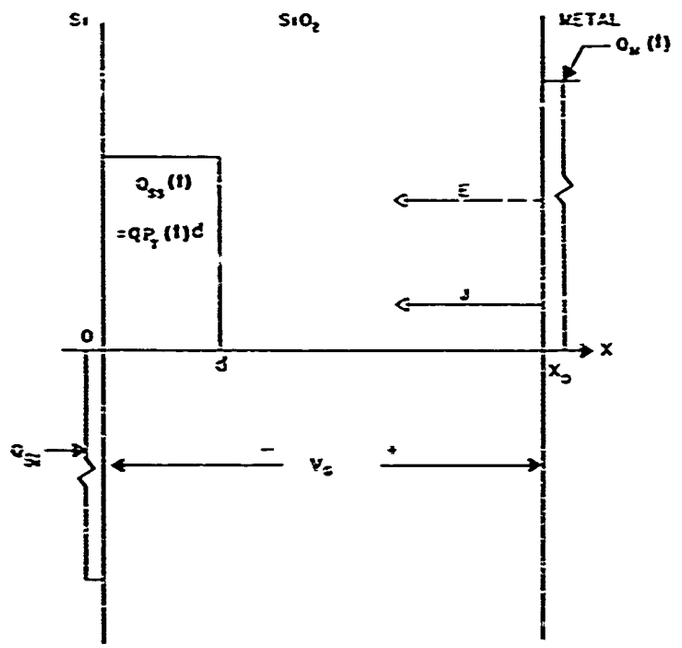


Figure 22. Model for Charge Buildup in SiO_2

Integrating and requiring $P_T(0) = 0$ gives

$$P_T(t) = \frac{V_G}{a}\left(1-e^{-abgt}\right) \tag{7}$$

Combining equations 2 and 7 gives

$$\Delta V(t) = -\alpha V_G\left(1-e^{-\beta gt}\right) \tag{8}$$

where $\alpha = \left(\frac{2x_o}{d} - 1\right)$ and $\beta = \frac{qd\mu\tau_e}{2\epsilon\epsilon_o x_o}$

If it is assumed that d is constant, then the expression for ΔV gives the correct behavior of charge buildup on dose, (gt), namely saturation. Furthermore, the dose rate, g, appears only in the product gt, and so the observed dose rate independence is correctly predicted. The saturation value of ΔV is αV_G. Experimentally $|\Delta V\ (sat)|$ is observed to increase linearly with V_G for low values of V_G (see Figure 7). At higher values $|\Delta V\ (sat)|$ increases at a rate somewhat less than linear. Such behavior is to be expected if the traps near the

interface saturate and cause d to increase. An increase in d causes a decrease in α and hence a decrease in the slope of the ΔV (sat) versus V_G curve.

For the case of negative bias applied during irradiation, the positive space charge will collect at the metal-SiO_2 interface. The problem will be symmetrical as far as the charge buildup process is concerned, but the space charge will have much less effect on the Si in this case. An analysis of the buildup for negative bias values similar to that given above for positive values yields

$$\Delta V(t) = V_G\left(1 - e^{-\beta g t}\right) \tag{9}$$

where $\beta = qd \# r_e \; 2 \epsilon \epsilon_o x_o$

as before.

$\Delta V(t)$ saturates with increasing dose with $\Delta V(sat) = V_G$. Reference to Figure 7 shows that ΔV (sat) obeys this relationship at least approximately for the RN-1030 device.

From equations 8 and 9

$$\frac{\Delta V(sat)_+}{\Delta V(sat)_-} = \alpha = \left(2x_o \, d - 1\right) \tag{10}$$

i.e., the ratio of the voltage shift at saturation for a given positive bias to that for the same bias but with opposite polarity is independent of V_G. Table 1 lists the values of $\Delta V(sat)$ for the RN-1030 device at 10^6 rads for positive and negative values of V_G and the corresponding ratios.

With the exception of the result for $|V_G| = 10$ volts, the ratio is reasonably constant as predicted by equation 10. For $V_G = 10$ volts $\Delta V(sat)$ appears to be too small, perhaps for the reasons already discussed. The average value of the first

Table 1

Values of $\Delta V (sat)_+ / \Delta V (sat)_-$ for RN-1030 MOS-FET

V_G volts	$-\Delta V$ (sat) volts	$\Delta V (sat)_+ / \Delta V (sat)_-$
+ 1.0	9.2	5.8
- 1.0	1.6	
+ 2.5	16.6	6.5
- 2.5	2.55	
+ 5.0	29.4	5.9
- 5.0	5.0	
+10.0	38.5	3.6
-10.0	10.1	

three ratios is 6.1. Substitution of this value in equation 10 yields $d/x_o = 0.28$, indicating that the space charge is confined to a region of the oxide close to one or other of the oxide surfaces.

Equations 8 and 9 have been fitted to the results for the RN-1030 MOS-FET at bias values of +5 and -5 volts. The best fit for -5 volts is

$$\Delta V = -30\left(1 - e^{-3D}\right)$$

and for -5 volts

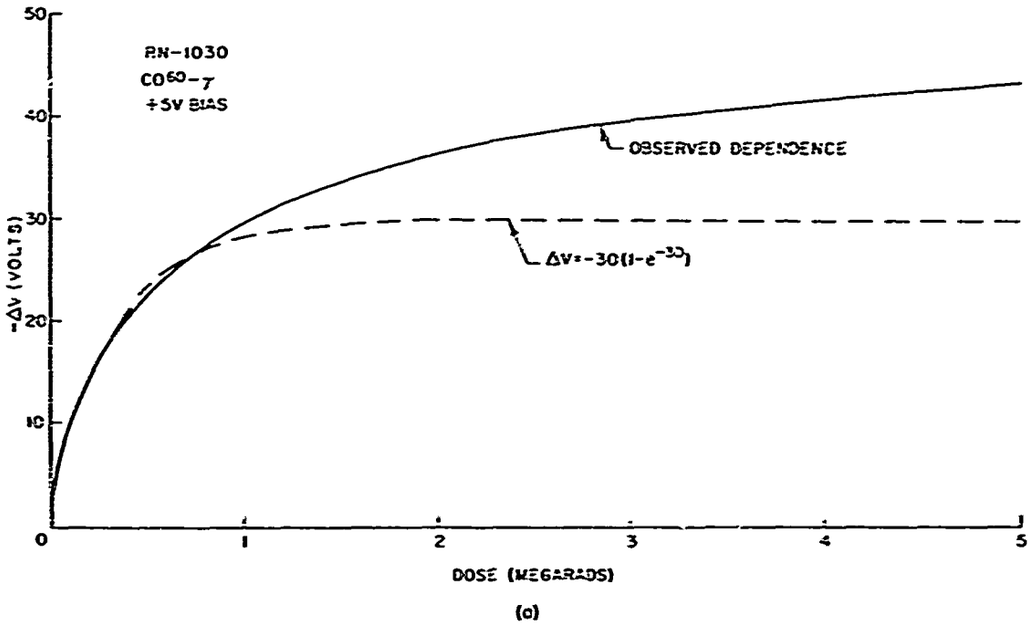
$$\Delta V = -5\left(1 - e^{-23D}\right)$$

where D is the dose in megarads. These curves are shown in Figure 23 (a) and 23 (b), together with the observed results.

For - 5 volts bias, the parameters α and β were chosen to give a good fit for lower dose levels. At higher dose levels, however, ΔV does not saturate in agreement with the calculated curve. The continued increase of ΔV may be caused by the effect of radiation on the "fast" surface states at the SiO_2-Si interface. It was observed that the C-V_G curves of irradiated devices showed a distortion which increased in magnitude with increasing dose. Furthermore, the distortion was considerably larger for positive values of gate bias. Distortion is usually interpreted in terms of fast surface states at the SiO_2-Si interface, and the increase of distortion may indicate an increase in the density of these states as a result of the Co^{60}-gamma irradiation. ΔV could continue to increase as long as the number of these states increases. For some reason which is not clear at present, the distortion does not appear to increase with dose as rapidly for negative bias values.

For -5 volts bias, the calculated and observed curves could be made to agree somewhat more closely. However, it was necessary to choose a much larger value of β (23 as against 3 for +5 volts bias).

If the oxide properties are uniform and the SiO_2-Si and SiO_2-metal interfaces are the same, then β should be the same for both bias polarities. There are two factors which may be at least partially responsible for the observed discrepancy. β depends linearly on the product $\mu\tau_e$ which was assumed constant throughout the oxide. However, since the SiO_2-Si interface region of the oxide is expected to have a higher concentration of imperfections than the rest of the oxide, both μ and τ_e may be significantly smaller in this region and thus produce a smaller value of β for bias positive values. The generation rate, g, may not be the same in both interface regions as was assumed in the development of the expressions for ΔV. (A variation in g will appear as a variation in β when the expressions are fitted to the experimental



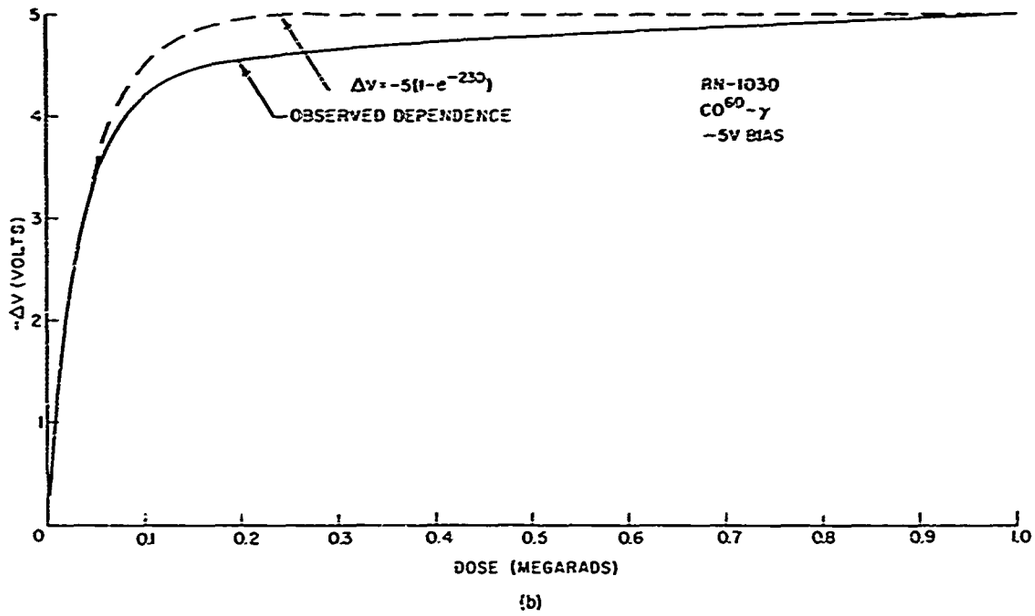


Figure 23. ΔV as a Function of Dose for an RN-1030 MOS-FET

curves) As the oxide structure changes, the energy required to form a hole-electron pair will change so that a given absorbed dose will produce more ionization in one part of the oxide than in another. Also, the fraction of absorbed energy which produces ionization may depend on the oxide structure. Furthermore, g in the interface regions may be influenced by the material adjacent to the oxide, i.e., metal for negative and Si for positive bias.

There appear to be a number of possible causes of the observed discrepancy in β (more accurately βg) for the two bias polarities. The controlling value of βg will be the value in the interface region where the space charge collects. The variation of βg across the oxide should not seriously affect the qualitative behavior of the space charge buildup.

The experimental results in Figure 23 (b) show that ΔV does not, in fact, saturate but continues to increase slowly in a manner similar to that at +5 volts bias. This observation suggests that the predicted curve should saturate at some value slightly less negative than -5 volts. In other words, if equation 9 is correct, the effective bias present across the oxide was actually smaller in magnitude than the 5 volts applied. Such would be the case if, for example, a small positive built-in field existed in the oxide. The fact that the minimum space charge buildup for the RN-1030 occurs at a small negative value of V_G (see Figure 7) also suggests that such a field exists.

If the equation for ΔV at + 5 volts bias and the doses indicated in Figure 23 (a) are taken as correct, then it is possible to estimate β and subsequently the product $\mu\tau_e$. If the energy required to produce one electron-hole pair is taken as 20 eV, then an absorbed dose of one rad produces $\sim 10^{13}$ pairs/cm^3. Using this fact and taking $\epsilon = 4$ and $d/x_o = 0.28$, then substitution in the expression for β yields $\mu\tau_e \sim 5 \times 10^{-16}$ m^2/V. For the RN-1030 MOS-FET's used in the experiments discussed earlier this value of $\mu\tau_e$ yields a ~ 10 in equation (1) and is, therefore, consistent with the assumptions made in deriving the expressions for ΔV.

The expressions derived for ΔV (equations 8 and 9) do not contain the temperature, T, explicitly. It is assumed that a hole once trapped will remain trapped. Elevated temperatures will, however, release holes from the traps and consequently reduce Q_{ss}. From the results of the experiments discussed earlier, it is apparent that the effect of temperature cannot be described in terms of a simple activation energy. Presumably, this result implies that the hole traps are not all located at the same energy level in the SiO_2 band gap. However, the effect of temperature during irradiation could be approximated by including a factor of the form $\exp(-E_T/kT)$ in the expressions for ΔV, where E_T is an average activation energy ~ 0.15 eV.

7. THERMOLUMINESCENCE OF IRRADIATED SILICON-DIOXIDE LAYERS ON SILICON SUBSTRATES

The SiO_2 passivation layer on planar semiconductor devices is sensitive to the effects of ionizing radiation. Radiation generally produces a positive charge buildup, Q_{ss}, in the SiO_2 near the SiO_2-Si interface. The present picture of the charge buildup process supposes that hole-electron pairs are created in the SiO_2 by the radiation; some of the pairs will, of course, recombine, but a fraction of the holes will be trapped while the corresponding electrons escape from the SiO_2 leaving a net positive space charge, Q_{ss}.

At present, no direct information regarding the nature of the assumed hole traps has been obtained. The standard methods of observing the buildup of Q_{ss} (e.g., capacitance-voltage curves of metal-oxide-semiconductor capacitors) do not yield any information about the trapping process itself. It is known, however, that Q_{ss} can be removed from the oxide by thermal annealing. The purpose of the work described in the following paragraphs was to determine whether luminescence accompanied the thermal annealing and, if so, the effects on the thermoluminescence of such factors as (a) type of oxide, (b) presence of impurities, particularly sodium, in the oxide, and (c) presence of an electric field in the oxide during irradiation.

a. Thermoluminescence Dosimeter

For preliminary investigations, a commercial thermoluminescence dosimeter manufactured by Controls for Radiation, Inc. was used. The instrument, under normal operation, uses a photomultiplier tube to observe the thermoluminescence from irradiated LiF samples. The electrical output from the photomultiplier tube is integrated and read on a digital scale.

For this experiment it was necessary to modify the instrument. To reduce the background infrared radiation, the photomultiplier tube was masked leaving a 3/8" square opening, the sample holder (planchet) was gold-plated and a blue filter placed between it and the photomultiplier tube. The shape of the sample holder made it necessary to use Si chips 3/8" square.

The output voltage of the photomultiplier tube was observed directly on an X-Y recorder, with the output of an Iron-Constantin thermocouple, embedded in the planchet, as the other input. The thermocouple was referenced to room temperature (25°C). Its output is roughly 0.055 mv/°C.

The silicon wafers were oxidized in three ways: (1) thermal growth in a dry oxygen atmosphere (dry), (2) growth in a wet oxygen atmosphere (steam), and (3) deposited by a silane process (deposited). The oxide layers were all approximately

5000 Å thick. Samples of each oxide also had a layer of sodium chloride deposited on the oxide surface. Sodium diffusion into the oxide was carried out at 600°C for 24 hours.

b. Procedure

Trapped charges (both holes and electrons) were created by irradiating the samples with Co^{60}-gamma rays. The magnitude of trapped charge was controlled by varying the radiation time from 20 minutes to 2 hours. After irradiation each sample was placed on the planchet and the thermoluminescence output plotted as a function of temperature.

For each irradiated sample three curves were generated (see Figure 24). Curve a was the photomultiplier output as a function of temperature for the empty planchet. This curve gives the real background. Curve b was the photomultiplier output of the sample and the planchet obtained from the first heating of the sample. It is this curve that gives the information concerning the amount of stored charge in the oxide layer. Curve c was generated under the same conditions as curve b, but for a second heating of the sample. This curve gives the effective background and indicates whether or not the trapped charges were completely removed by just one annealing.

An experiment was carried out to determine whether room temperature annealing in the time interval between removal from the radiation cell and measurement of light output significantly affected the light output. A sample was irradiated

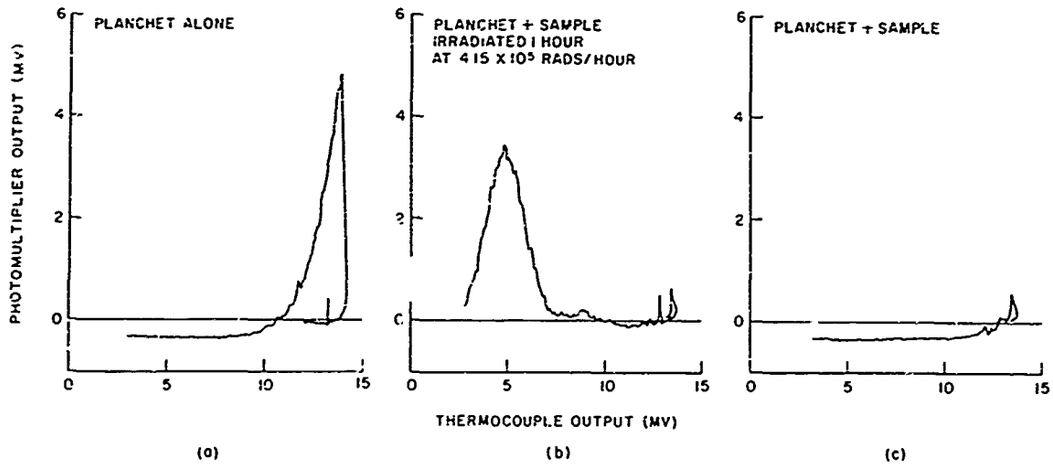


Figure 24. Photomultiplier Output of Sample and Planchet as a Function of Thermocouple Output

for 40 minutes, and then allowed to remain at room temperature for a specific time period. At the end of this time interval the luminescence was observed. As can be seen from Figure 25, there is a significant reduction in the light output for periods $\geq$ 1 hour, but for periods of $\leq$ 20 minutes the output was reduced by $\leq$ 5%. Therefore, for this experiment all measurements were taken within 20 minutes of the time the sample was removed from the Co^{60}-gamma cell.

In order to compare different oxide specimens, it was first necessary to determine their relative luminescence efficiency. Luminescence efficiency is effectively the fraction of the total energy released by annealing which appears as luminescence. The samples were placed in an electron gun and mounted in such a way that the light emitted by the oxide was observable through a window in the gun chamber. The samples were then bombarbed with 20 keV electrons and the light output from several samples of different oxide types was compared by several observers. It was agreed that there was no difference detectable to the unaided eye between the light output from various samples. Thus, it was concluded that the efficiency was the same for all oxide types within the accuracy required for these experiments. As a result, any differences seen in the light output for different oxides due to thermoluminescence are not caused by differences in the luminescence efficiencies of the oxides, but by differences in the amount of stored charge Q_{ss}.

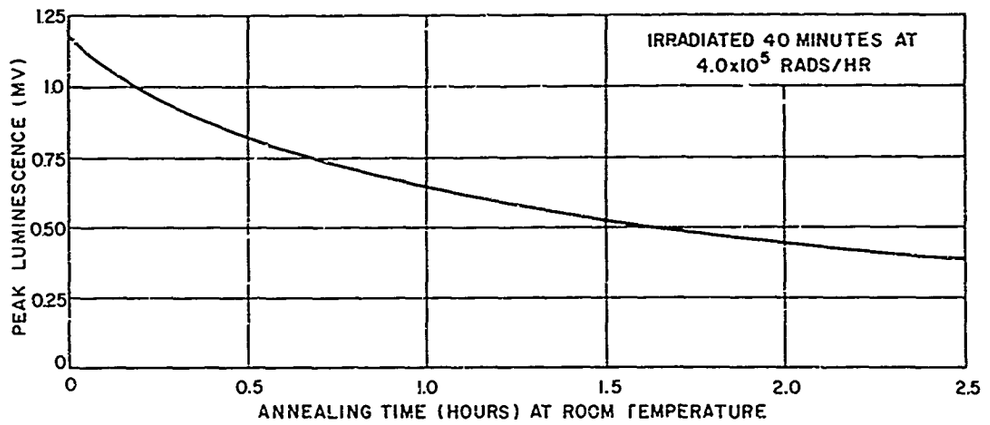


Figure 25. Photomultiplier Output of Contaminated Steam Grown Oxide as a Function of Room Temperature Annealing

c. Factors Affecting Thermoluminescence

The type of oxide was the first factor studied which might affect the thermoluminescence. The oxides were irradiated with Co^{60}-gamma rays for various lengths of time and then placed in the thermoluminescence dosimeter. A plot of photomultiplier output as a function of thermocouple output was generated for each oxide. Typical results for each type are shown in Figure 26.

A large thermal anneal peak is observed for each oxide, varying from 85 C in the dry oxide to 120 C in the deposited oxide. In addition, there is some evidence of a subsidiary peak between 160 and 190 C. The principal peak corresponds to that seen in the thermal anneal of irradiated MOS-FET devices made with a deposited oxide layer.

Q_{SS} will be proportional to the integrated light intensity, i.e., the total area under each of the curves in Figure 26. The areas under these curves have been determined as a function of integrated dose. The results are given in Figure 27. In all cases the trapped charge appears to saturate in the neighborhood of 1 Mrad. This is also the dose at which Q_{SS} saturates in MOS-FET devices. It is evident that the integrated luminescent output and presumably the magnitude of stored charge Q_{SS} is ≈75% greater in the steam (wet) grown oxide than in the other two types.

The next factor studied was the effect Na contamination had on the luminescence of the different oxides. The same irradiation and measurement cycle used for the uncontaminated oxides was employed. Typical curves for the Na-contaminated

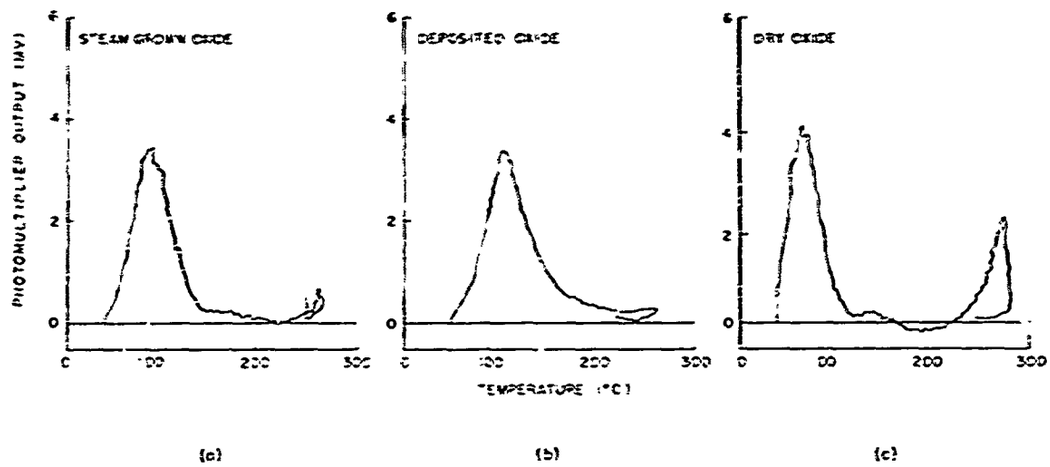


Figure 26. Photomultiplier Output of Three Types of Uncontaminated Oxides as a Function of Temperature (4.15×10^6 rads)

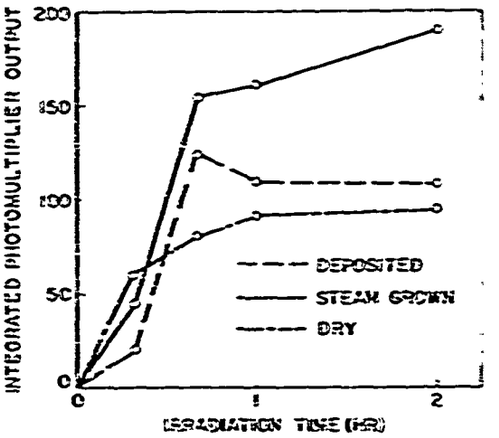


Figure 27. Integrated Output of the Three Types of Uncontaminated Oxides as a Function of Irradiation Time

oxides are shown in Figure 28. The introduction of sodium shifts the peak of the annealing curves to lower temperatures to 65°C in the dry oxide, and to 100°C for both steam (wet) and deposited oxides. In steam and deposited oxides a distribution of high temperature (deep) traps also appears to have been added. These would cause the broad luminescence at temperatures above the principal peak. Again, in order to compare the oxides, the area under the curves must be compared. Figure 29 shows that the presence of sodium affects the luminescence of each oxide in a

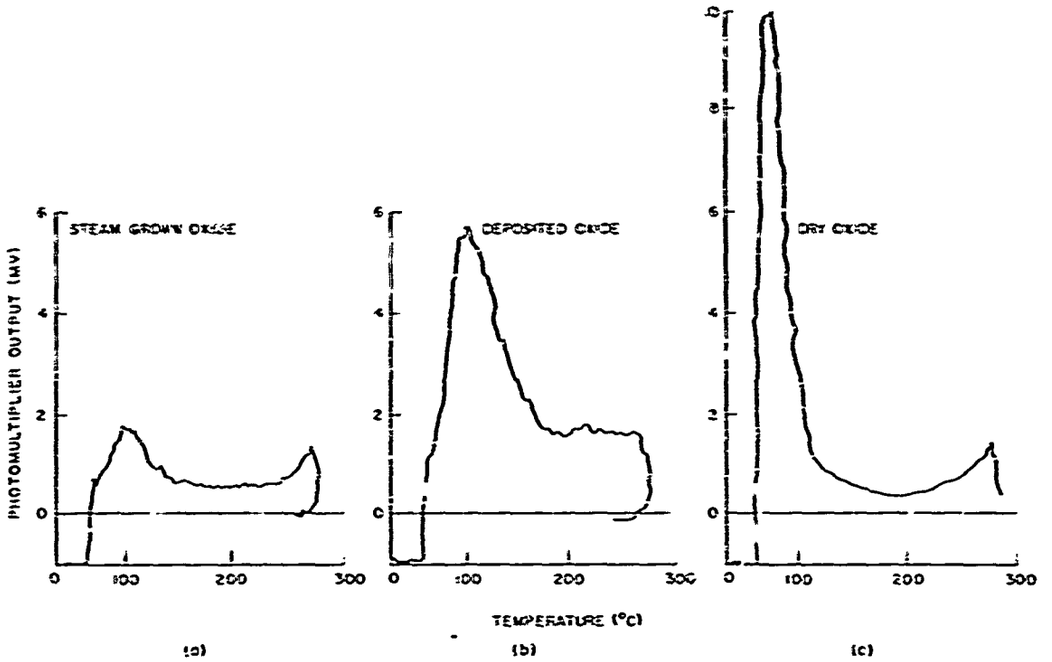


Figure 28. Photomultiplier Output for Three Types of Contaminated Oxides as a Function of Temperature (4.15×10^6 rads)

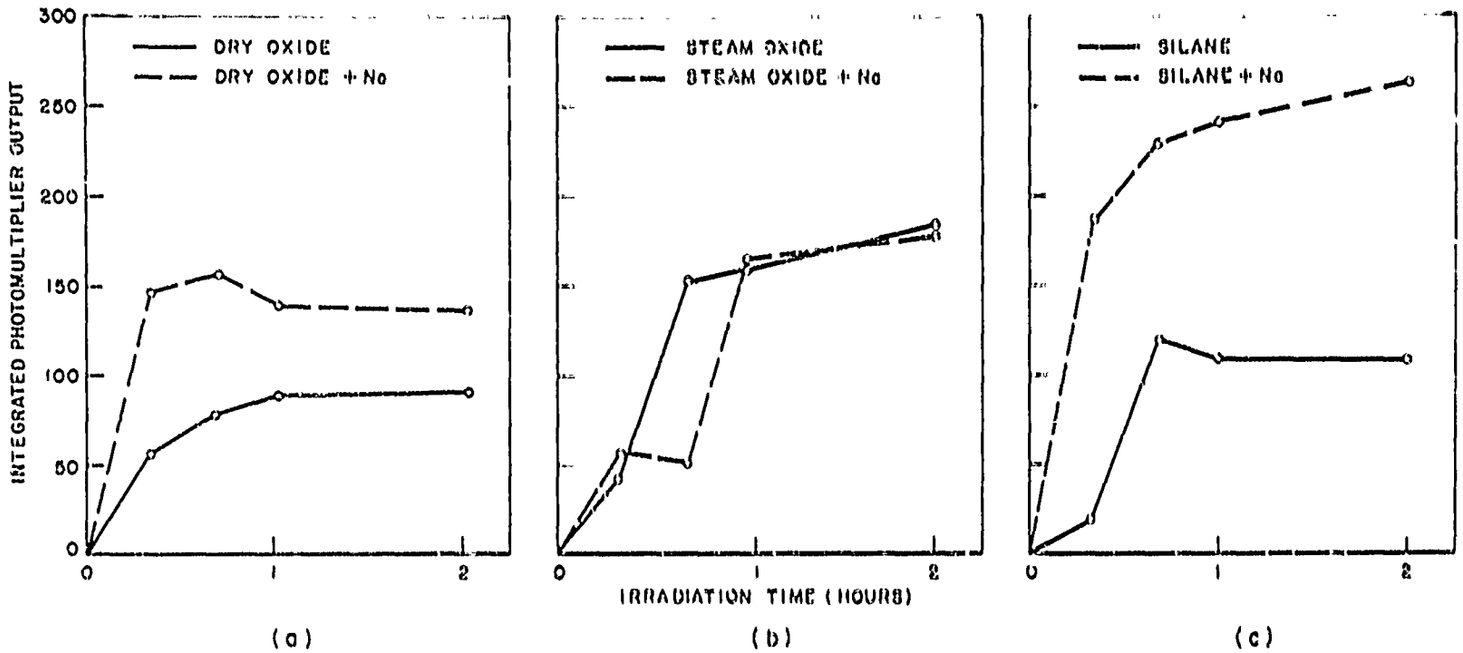


Figure 29. Integrated Output for Uncontaminated and Contaminated Oxides as Functions of Irradiation Time

different way. There was essentially no change in the luminescence for the steam grown oxide. However, sodium increased the luminescence from the dry oxide by about 50%, and from the deposited oxide by approximately 250%.

The presence of an electric field in the oxide during irradiation was the last factor studied. The electric field was applied across the oxide by means of a ball of mercury on the surface of the oxide and a gold wire bonded to the silicon substrate.

Uncontaminated oxides were irradiated for one hour at different biases, ranging from -30V to +30V. The results of this study are shown in Figure 30. The bias caused only a slight effect on the luminescence, except at -30V where a reduction of about 30% was observed. This result is indeed surprising since other work has shown a very strong dependence of the charge accumulation on bias.

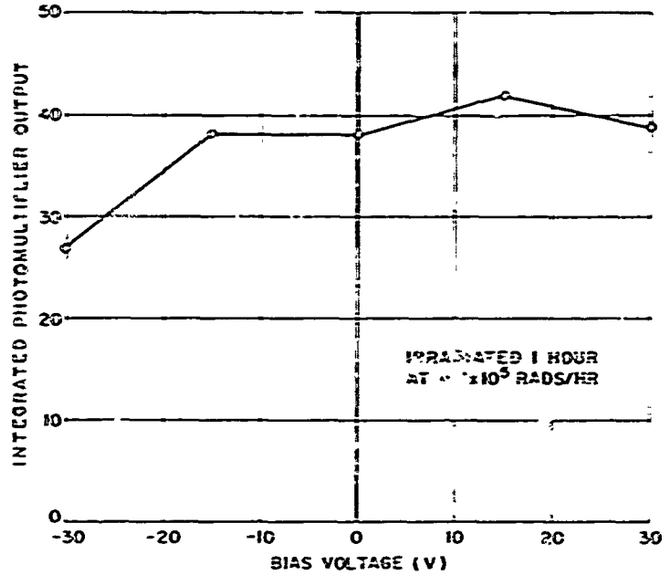


Figure 30. Integrated Luminescence of Uncontaminated Steam Grown Oxide as a Function of Bias Voltage

d. Conclusions

The thermoluminescence technique would seem to be quite useful in studying the problems associated with trapped charge in oxidized silicon surfaces. It is evidently a sufficiently sensitive tool to detect differences in the process used to prepare the oxide. Consequently, this technique could be used routinely to check oxidized wafers for contamination introduced either during oxide growth or in subsequent device processes. The technique should also be of use in selecting oxides (or other insulating surfaces) to be used in an ionizing environment.

Although it has not been possible to demonstrate that the traps which store charge next to the SiO_2-Si interface in oxidized devices are related to those observed in thermoluminescence, it would not be a difficult task to prepare a number of MOS capacitors on the same slices that are used for thermoluminescence studies to establish the necessary correlation. Some correlation exists in that the thermal anneal peak in irradiated MOS-FET's corresponds to the thermoluminescence peak seen in the comparable oxide.

Trap depths (E_t) giving rise to the thermoluminescence can be estimated using the expression

$$E_t = k \frac{T^2}{\Delta T}$$

where T is the temperature of the peak, ΔT is the temperature difference between half-maxima, and k is the Boltzman constant.

Using this expression, trap depths are estimated at 0.25 eV in steam oxide, 0.30 eV in dry oxide, and 0.22 eV in deposited oxide. The latter value should be compared with the value 0.15 eV obtained for hole traps responsible for the trapped charge Q_{SS} in irradiated MOS-FET's using this oxide. Introduction of sodium into these oxides apparently introduces still shallower traps estimated to have a trap depth of 0.2 eV. This result indicates that sodium was probably not the impurity that caused the trapping observed in uncontaminated samples.

In the case of the oxidized surface to which no metal contact is made, it must be assumed that irradiation causes both holes and electrons to be trapped. In contrast, the electrons in a MOS structure are presumed to move to the metal electrode and leave the oxide. Since the shallower of the two traps, hole and electron, would control the annealing behavior, it is concluded from the agreement observed between hole trap level in MOS-FET's and the thermoluminescent trap level that the thermoluminescent traps are also hole traps.

The failure to observe any marked change in the luminescence when bias was applied during irradiation is surprising, since a marked dependence on bias is observed for the number of holes trapped at the SiO_2-Si interface in irradiated MOS

devices. It may be that bias does not alter the total trapped charge in either case, but only alters the relative disposition of trapped holes and electrons. This feature merits further investigation.

The thermoluminescence technique appears to have some very real value for studying trapping phenomena in oxidized layers. Furthermore, a correlation of such studies and parallel studies on MOS capacitors could provide some real insight into both the radiation damage and reliability problems in oxidized silicon surfaces. Such correlated studies could also prove useful in unraveling the mysteries of trapping phenomena, both latent and radiation-induced, in a wide variety of insulating materials.

8. EFFECTS OF Co^{60} - GAMMA IRRADIATION ON SCHOTTKY BARRIER DIODES

a. Introduction

A study was carried out of the effects of gamma irradiation on Au-Si and Cr-Pa-Pt-Si Schottky barrier diodes. In theory, such devices — being majority carrier structures — should be relatively insensitive to radiation damage. However the I-V characteristics of a Schottky barrier are closely related to interface states at the semiconductor-metal junction. In this respect, the Schottky barrier may be similar to the MOS structure with an extremely thin "O" layer, so that the Fermi level of the interface states is closely coupled to the Fermi level in the metal. In a thick "O" layer, the interface Fermi levels couple to the semiconductor. High densities of interface states (10^{13} states cm^2) have been found for many metal-semiconductor junctions.[11] These interface states explain the lack of strong dependence of barrier height on the metal work function.

Energetic radiation could be expected to change the interface state density and thus the I-V characteristics. However very large doses of Co^{60} -gamma radiation ($>10^8$ rads) would be required to affect significantly barriers whose initial interface state densities were of the order of 10^{13} states/cm^2.

b. Experimental Procedure

The devices studied were WECo 479A Au-Si and developmental Cr-Pa-Pt-Si diodes having roughly the same structure. This structure is shown in Figure 31. Pre- and post-exposure measurements of I-V characteristics at -50°C, 25°C, and 100°C were made with a logarithmic curve tracer which will be described later in this report. Typical I-V characteristics before and after exposure are shown at

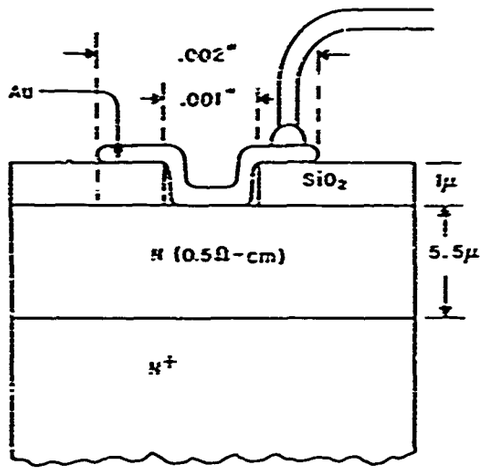


Figure 31. Structure of Au-Si Diodes

the three temperatures for a Cr-Pa-Pt-Si diode in Figures 32 and 33. Capacitance-voltage measurements were made with a Boonton Capacitance Bridge. Exposures were made in a Co^{60} gamma cell at a rate of 0.4×10^6 rads/hr. Devices were exposed with and without bias. Several devices were also exposed to fast neutrons from the White Sands Fast Burst Reactor. Both gamma and neutron irradiated devices were subsequently subjected to a series of isochronal anneals.

c. Results

A wide variety of responses to ionizing radiation were observed for the I-V characteristics. The Au-Si production devices exhibited a narrower range of responses than did the Cr-Pa-Pt-Si developmental devices. For example reverse current changes at -5 V ranged from virtually zero to nearly four orders of magnitude increase at 10^8 rads integrated dose. Figures 34 and 35 show the average increases in reverse current at -5 V for both groups of devices as well as the best and worst device in each group. There is evidence of saturation in the damage mechanism at integrated doses of 10^8 rads. These large and variable responses were evident also in the forward characteristic up to ≈ 100 ma. Above this current, relatively small changes in the forward characteristic were observed. C-V characteristics show a $1/C^2$ vs. V characteristic with an extrapolated barrier height of ~0.40 V for Au-Si and 0.36 V for Cr-Pa-Pt-Si devices. Essentially no change in this extrapolated barrier height occurred for integrated doses up to 10^8 rads. This result indicates that changes in interface state density are small and do not contribute to the observed changes in the I-V characteristic. The devices subjected to energetic neutron bombardments showed negligible changes for integrated fluxes up to 10^{14} N/cm^2. Since this neutron fluence should produce approximately the same lifetime degradation as the largest gamma dose employed in this study, it is concluded that changes in minority carrier lifetime do not significantly contribute to the observed changes in I-V characteristics.

Because of large variations among devices, there was no effect ascribable to applied bias during test. Furthermore, although the variety of responses to ionizing radiation was less for the Au-Si devices, there were no other obvious differences in the response to radiation of the two types of contacts.

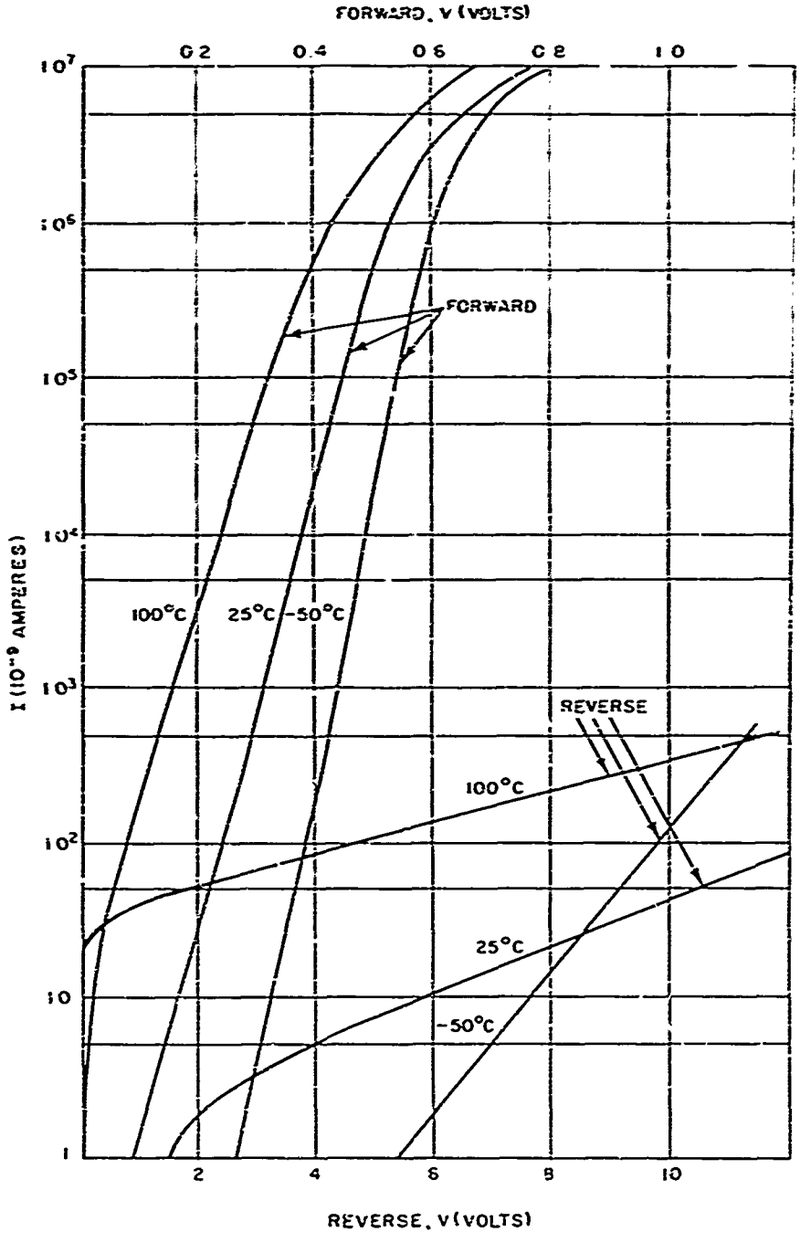


Figure 32. Current-Voltage Characteristics for a Cr-Pa-Pt-Si Diode Before Exposure

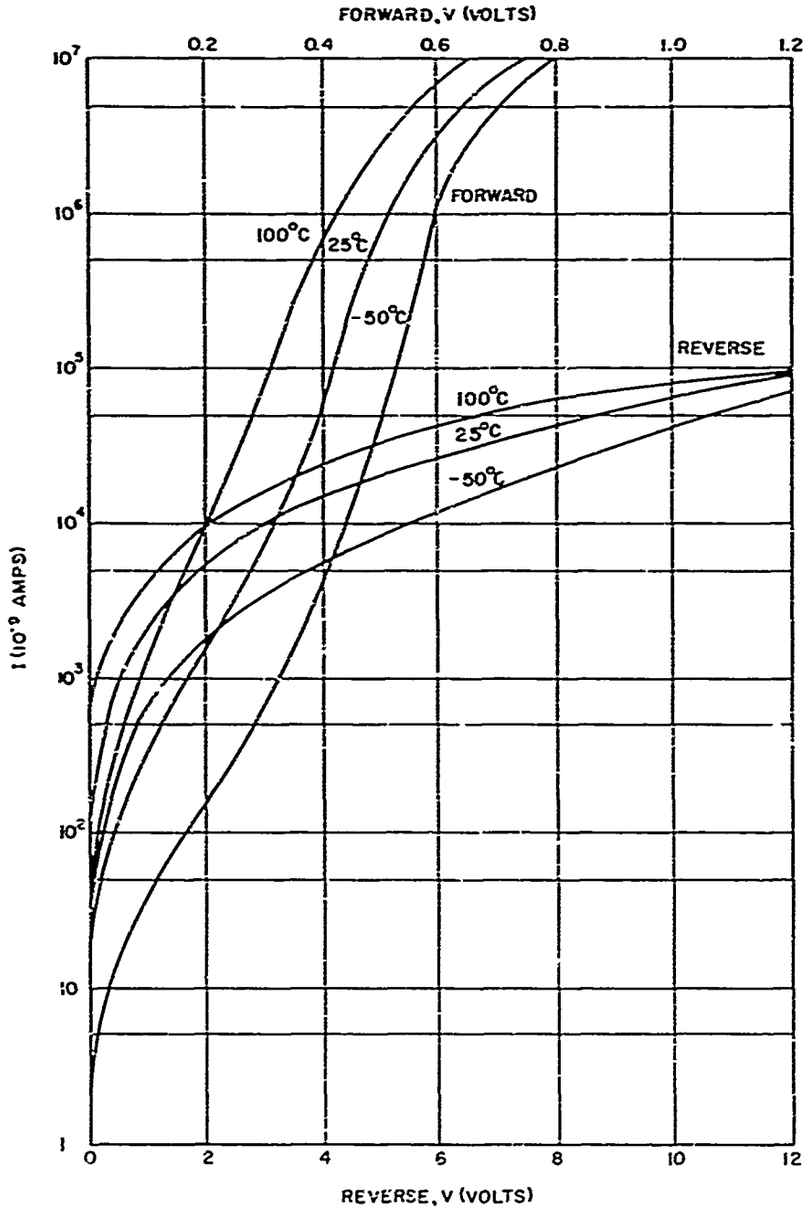


Figure 33. Current-Voltage Characteristics for a Cr-Pa-Pt-Si Diode After Exposure to 10^8 rads

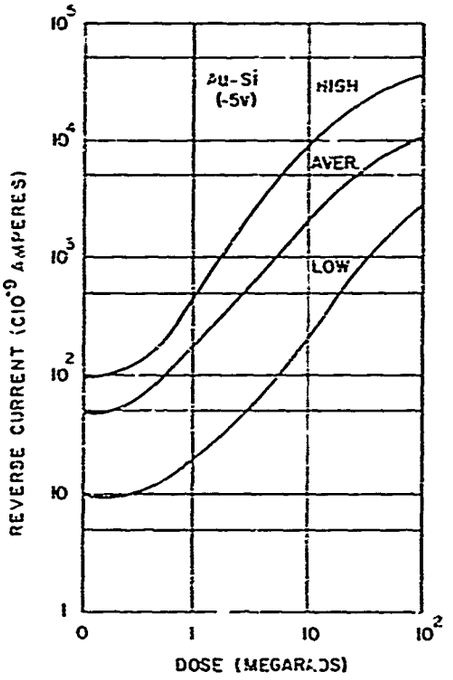


Figure 34. Increase in Reverse Current at -5 Volts for Au-Si Diodes

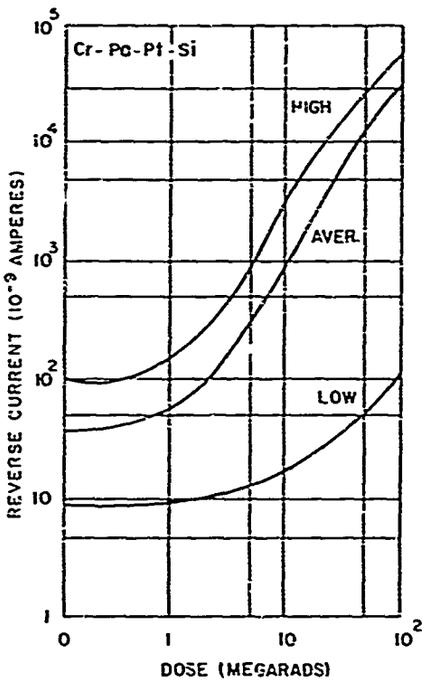


Figure 35. Increase in Reverse Current at -5 Volts for Cr-Pa-Pt-Si Diodes

Those devices which had very large increases in reverse current with radiation showed a much reduced dependence of reverse current on temperature. For example, devices with reverse currents at -5 V of 20 - 50 μa showed a factor of 2 change in reverse current for a temperature change of 75°C in contrast to the order of magnitude change observed in devices having reverse currents $<$ 1 μa at -5 V. Annealing of radiation damage in devices showing large changes in reverse current was significant for temperatures below 150°C. Annealing in devices with small increases in reverse current required temperatures of the order of 300°C.

A correlation was observed between the forward spreading resistance of these devices and the magnitude of increase in reverse current with irradiation. The correlation is evident in Figure 36 for the Cr-Pa-Pt-Si devices after 10^8 rads irradiation. Forward spreading resistance is determined from the slope of the I-V characteristic for forward currents between 1 and 10 ma.

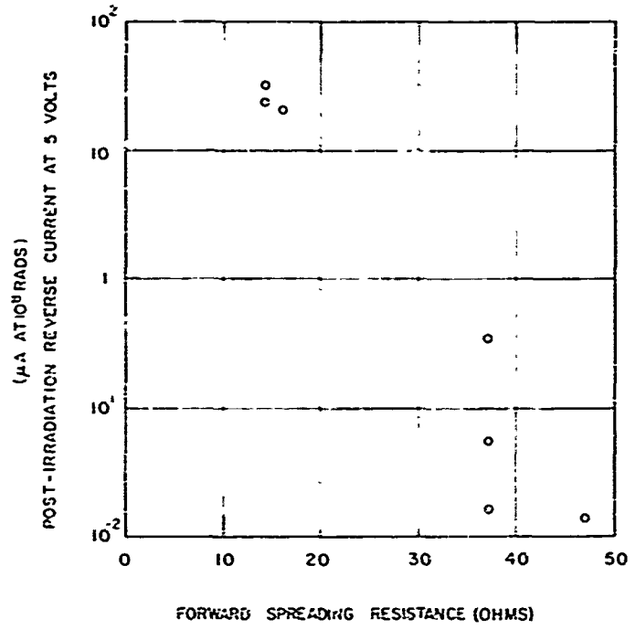


Figure 36. Forward-Spreading Resistance versus Reverse Current for Cr-Pa-Pt-Si Diodes

d. Model for Radiation Damage in Schottky Barrier Diodes

A model has been proposed for the observed Co^{60} -gamma radiation damage in Schottky diodes for those devices which show very large increases in reverse current. This type of damage saturates in the neighborhood of 10^8 rads and anneals out at relatively low temperatures (150°C or less). These observations indicate that this type of damage is related to the charge storage process in the oxide layers of MOS devices, i.e., to the buildup of trapped positive charge at the SiO_2-Si interface. This charge would develop an enhancement layer on the n-type silicon epitaxial layer. This enhancement layer could affect the I-V characteristic only where the metal, SiO_2 and silicon have a common boundary, i.e., at the perimeter of the metal-semiconductor contact. At this perimeter the metal is in contact not only with the moderate resistivity epitaxial silicon but also with a high conductivity surface enhancement region. The metal-semiconductor depletion layer will be drastically narrowed in the enhancement region. In addition the radius of curvature at the edge of the depletion region is reduced. Both of these effects, which are illustrated in Figure 37, will drastically reduce the breakdown voltage between the metal and semiconductor. Because of the charges present in the oxide, a large field is produced in the silicon without application of reverse

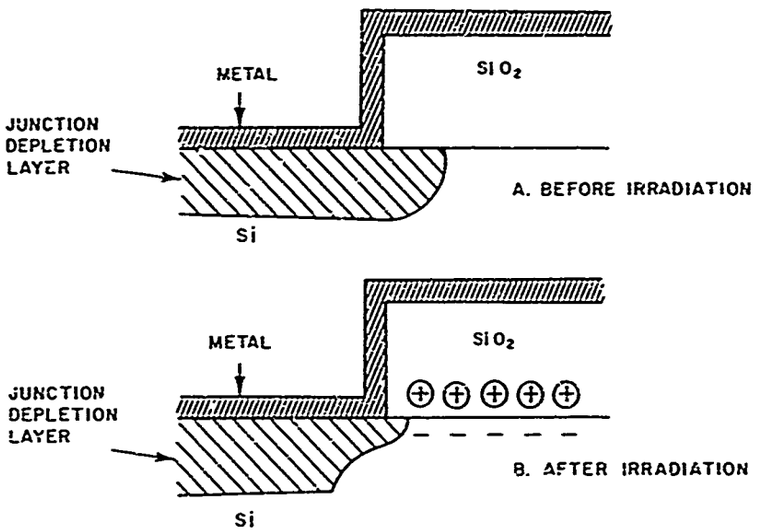


Figure 37. Effect of Radiation on Schottky Barrier Diodes

bias. Leistiko and Grove[12] have pointed out that for surface charges in excess of $10^{12}/cm^2$, surface breakdown will take over in typical silicon planar devices. Grove and Snow[7] have reported surface charge densities for oxidized surfaces exposed to ionizing radiation in excess of 10^{13} cm^2, so that surface induced breakdown should occur. Ultimately for very strongly accumulated surfaces tunnel current will flow from metal to semiconductor. Both breakdown and tunnel currents would be expected to show a small temperature dependence as was observed in the irradiated diodes having very large reverse currents.

In the production of these Schottky barrier diodes, it is observed that the surface metalization sometimes fails to sinter to the silicon next to the oxide so that only a fraction of the total metal-semiconductor perimeter adjoins oxide. Because of the reduced metal-semiconductor area the forward spreading resistance is increased in these devices. Thus the correlation evident in Figure 36 between reverse current after irradiation and forward spreading resistance is explained by the model.

e. Summary

A variety of responses to ionizing radiation have been observed in the I-V characteristics of Au-Si and Cr-Pa-Pt-Si Schottky barrier diodes. Changes in barrier height or minority carrier lifetime do not account for the results.

Surface charge buildup in the oxide layer adjacent to the metal-semiconductor perimeter will lead to surface breakdown and possible metal-semiconductor tunneling and probably accounts for the large increases in reverse current. Devices with incomplete metal-semiconductor contacts will have less perimeter in common with oxide and show less radiation sensitivity. Reduced sensitivity to radiation effects in Schottky barrier devices of the type studied here (planar) will require development of oxide (or insulator) isolation layers having significantly lower ionizable trap concentrations.

III. DEVICE STUDIES

9. INTRODUCTION

The principal effort in device studies has been concentrated on bulk effects in silicon devices caused by fast neutrons. Recovery effects in transistors and diodes after a single pulse of fast neutrons have been studied with a view toward understanding the nature of neutron damage centers. The observed recovery effects are believed to be caused by multi-charged damage clusters, which act as traps rather than recombination centers.

The effects of neutron damage on high-injection-level operation in transistors have been examined.* This study revealed basic limitations in ultimate power-switching capabilities of transistors which can be overcome by using PNPN devices. The response of narrow-base PNPN devices to neutron damage has also been studied and compared to a fairly simple device theory. Design considerations for PNPN devices to be used in radiation environments have been developed.

10. SHORT-TIME RECOVERY FROM NEUTRON DAMAGE IN HIGH-FREQUENCY SILICON DEVICES

Rapid recovery of d-c gain at room temperature in silicon transistors irradiated with bursts of fast neutrons has been reported by Sander and Gregory[1] and by Binder and Butcher.[2] These authors reported that significant recovery of current gain occurred at times of the order of milliseconds to seconds. Such recovery could drastically affect the performance of missile electronics within seconds after exposure to a weapon environment.

To shed further light on the matter, a variety of high-frequency narrow-base silicon transistors and one type of PIN diode were exposed to pulses of neutron from the White Sands Fast Burst Reactor. Base current (i.e., current gain) in the

*Some of the results published herein, particularly Attachment II by E.A. Overstreet and part of the work reported in Attachment III by D.K. Wilson and H.S. Lee, were supported by Contract DA-30-069-AMC-333(Y). However, because of their relevance to AF 19(628)-4157 studies they have been included.

transistors and forward voltage drop in the PINs were monitored before, during, and after the burst. The effects of various bias conditions and of temperature on both transistors and diodes were explored.

The results indicate that the ratio of recoverable neutron-induced recombination rate to the stable neutron-induced recombination rate, R_A/R_N, is of the order of unity at times of the order of 1 msec. Thus, the current gain of a transistor immediately following a burst of neutrons might be as little as one half of its final gain for times of the order of 1 msec. R_A/R_N was always highest at low emitter or forward currents.

The rates of recovery were observed to be strongly dependent on injection level in the devices studied. Little or no recovery occurred at times up to 0.5 sec without minority carrier injection.

The ratio, R_A/R_N, at 1 msec after burst was observed to be temperature dependent, increasing to greater than two at temperatures of -50°C. However the rate of recovery of R_A/R_N was relatively insensitive to temperature.

Two additional effects were observed in neutron-irradiated devices, and they appear to be related to the recovery process. The first was observed in neutron-irradiated devices which were heated to anneal some of the neutron damage. When these devices are returned to room temperature a slow reverse recovery, current gain decreasing with time, was noted. The second effect was observed in capacitance measurements of reverse-biased silicon diodes. After a burst of neutrons, a large decrease in capacitance was measured which recovered very slowly to the final measured capacitance change. Forward bias accelerated the capacitance-recovery and reverse-recovery processes markedly.

The recovery effect may be important in the case of devices whose performance is marginal in a weapon environment. Such devices include high-power transistors, PNPN switches, and PIN diodes. Since recovery is slow when minority carriers are not injected, device degradation will be most evident in circuits having a very low duty cycle.

A model is proposed for the observed behavior which attributes recovery to changes in the charge state of the recombination and trapping centers, i.e., to electronic processes. This model is compared with that of Sander and Gregory, which assumes that atomic migration of the recombination defects occurs.

a. Experimental Procedure

The devices were exposed to a short burst (approximately 100 μsec long) of fast neutrons (E > 10 kev) with a fission spectrum distribution from the White Sands Fast Burst Reactor. The neutron fluence received by each device was of the order of 1 to 5 x 10^{13} N/cm^2 as determined by sulphur foil activation methods. Unless otherwise noted, device temperatures were approximately 30°C, except for the medium- and high-power devices operated at collector currents greater than 100 ma. The transistors were supplied with constant emitter current and the diodes with constant forward current during measurement. Current gain in the transistors was determined by measuring the voltage drop produced across a known resistance by the base current. Voltage changes with time were recorded using a Minneapolis Honeywell Visicorder with galvanometers having a response time of less than 1 msec. The circuit used for this measurement is shown in Figure 38. Forward voltage drop across the PIN's was also measured with the MH Visicorder and the circuit shown in Figure 39. Only the devices under test and the cables to supply potentials were exposed to the immediate environment of the reactor. Ferrite cores were placed around the emitter leads of the transistors to eliminate high-frequency singing. In both circuits provision was also made for delaying the application of emitter current or forward current to the devices for times from 1 msec to 1/2 sec after the radiation burst. The delayed bias circuit could also be operated under a variety of bias conditions during the reactor pulse.

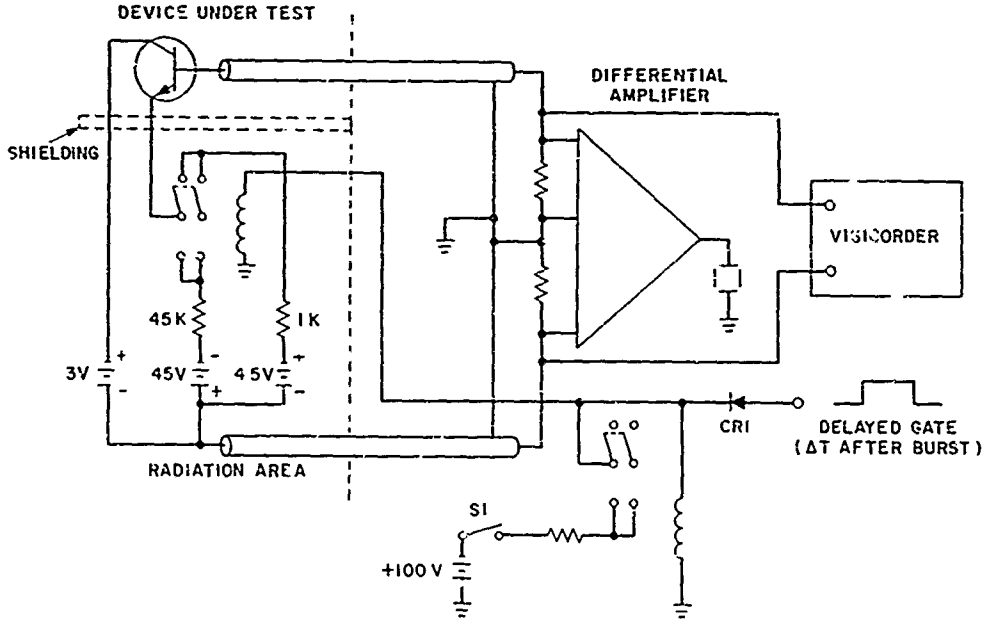


Figure 38. Transistor Fast Anneal Test Circuit (for I_c = 1 ma)

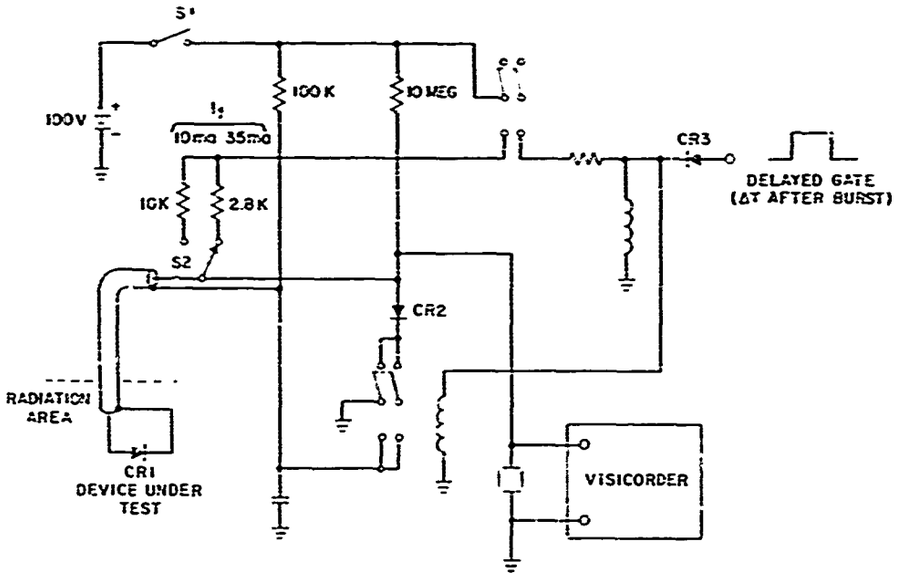


Figure 39. PIN Fast Anneal Test Circuit

Devices were also exposed at various temperatures using the hot cold mount shown in Figure 40. Low temperatures were obtained with this mount by passage of cold nitrogen gas through the transistor support and high temperatures were obtained by heaters embedded within the mount. A chromel-alumel thermocouple referenced to ice-water temperature was used to monitor the device temperatures during exposure. In the results cited here, temperatures from -50 C to +100 C were employed.

b. Results

(1) Transistors. All of the transistors exposed in these studies were planar and had basewidths of the order of 1μ except for the WECo 2N-560 (W = 3μ). The planar types included low power NPN's (WECo A-2409) and PNP's (Motorola MM999), medium power NPN's (Motorola 2N-3252) and PNP's (Motorola 2N-3245), and a high power NPN (Motorola SF2585A). Typical dc current gains before neutron bursts of 5 x 10^{13} N/cm^2, as well as several hours after, (for a variety of collector currents) for these devices are shown in Table 1.

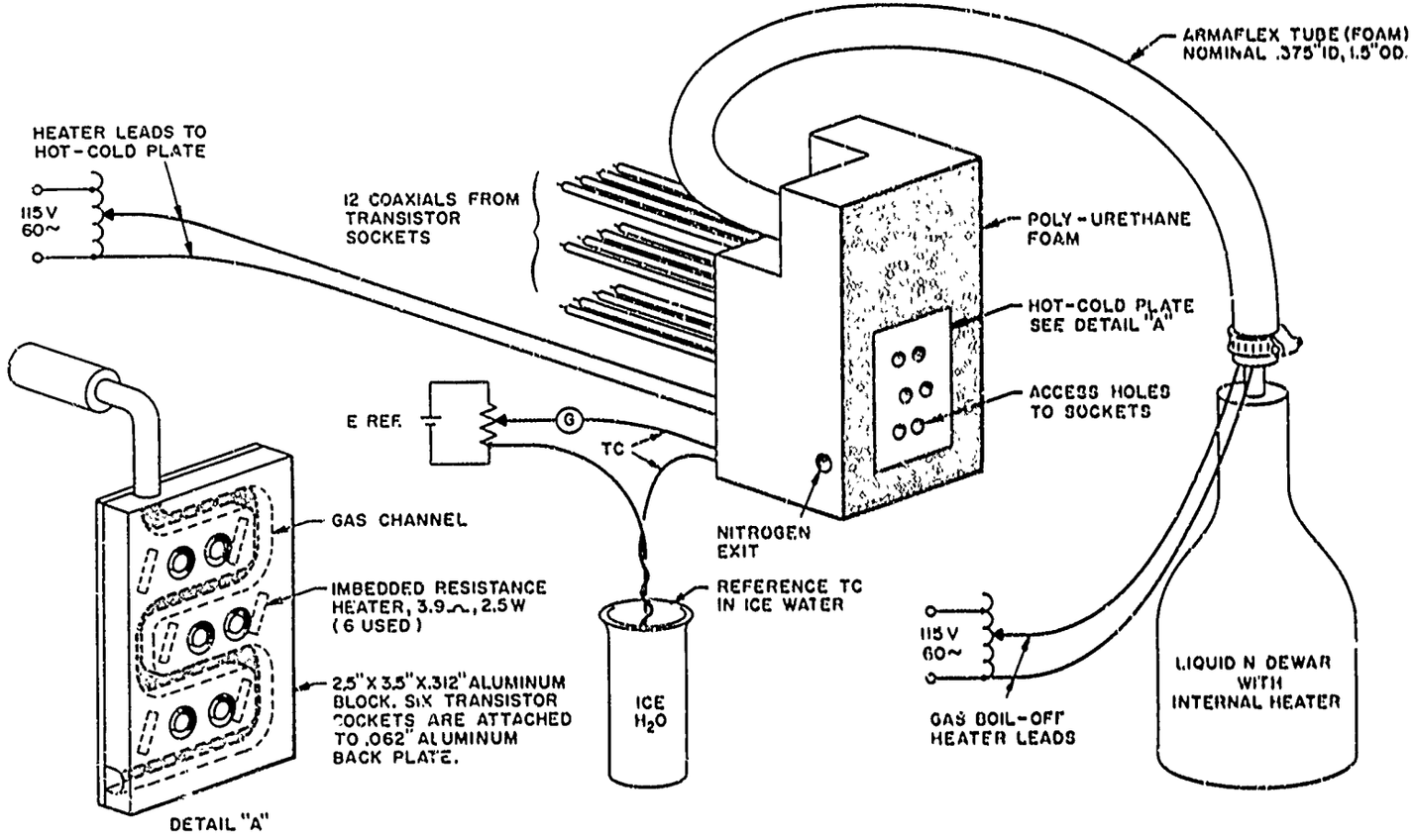


Figure 40. Hot-Cold Mount for Transistors and Diodes

TABLE 1

Typical D-C Gains of Experimental Devices
Before and After Exposure to 5×10^{13} N/cm^2

Device	Collector Current (ma)	h_{FE} Before	h_{FE} After
2N560	1	24	8
2N560	10	40	12
A2409	1	69	26
A2409	10	83	40
MM999	1	124	22
MM999	10	126	36
2N3252	1	67	9
2N3252	100	92	27
2N3244	1	95	11
2N3244	100	130	37
SF2585A	10	100	17
SF2585A	400	100	27

Collector voltages during most of the test were 3.0 volts for all devices, except for the SF2585A where a collector bias of 1.5 volts was used. Current densities in the transistors for 10 ma were of the order of 100 amp/cm^2 for the low-power NPN's and PNP's, 10 amp/cm^2 for the medium-power NPN and PNP's, and 1 amp/cm^2 for the high-power NPN's.

Figures 41 through 45 show the recovery of d-c gain normalized to the final current gain for these devices after a burst of 5×10^{13} N/cm^2. In all cases transistor current gain was at a minimum immediately after the neutron burst and recovered subsequently to a higher final value. Where several devices of the same type were studied, variations among devices were small. Recovery times were generally comparable for all of the devices studied including both NPN and PNP types. It will be observed that in all cases the magnitude of the annealing was larger at low collector currents. In the delay bias arrangement it was determined that recovery did not occur until forward emitter current was passed. This was true for delays in application of emitter current up to 1/2 sec after the reactor pulse. In most cases the same magnitude and time of recovery were observed for delayed application of bias as were observed when forward bias was applied during the reactor pulse.

If it is assumed that neutron damage results in an increased recombination rate due to damage centers in either the emitter space charge region or the neutral

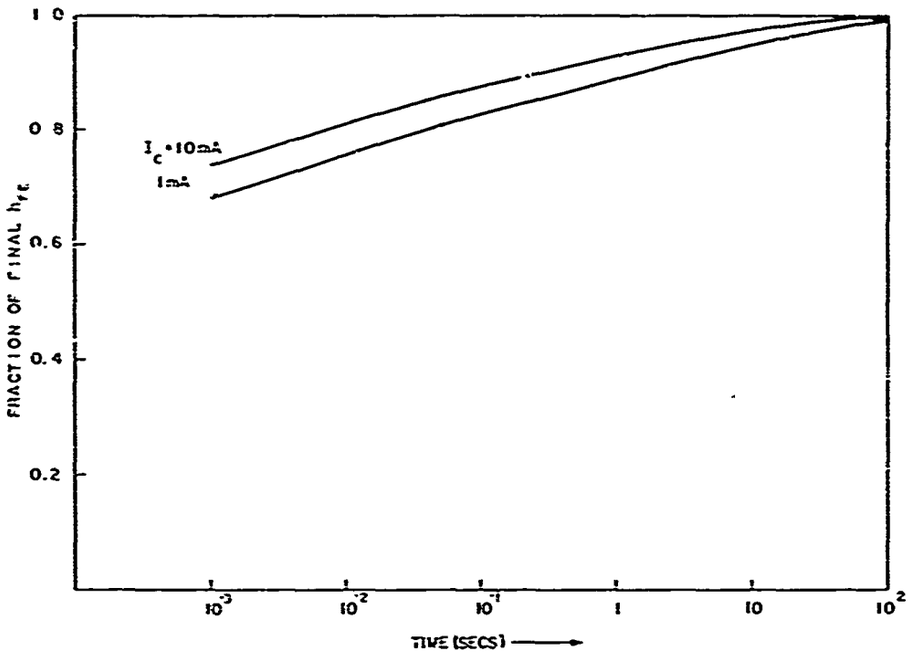


Figure 41. Recovery of 2N560 After 5 x 10^{13} N/cm^2

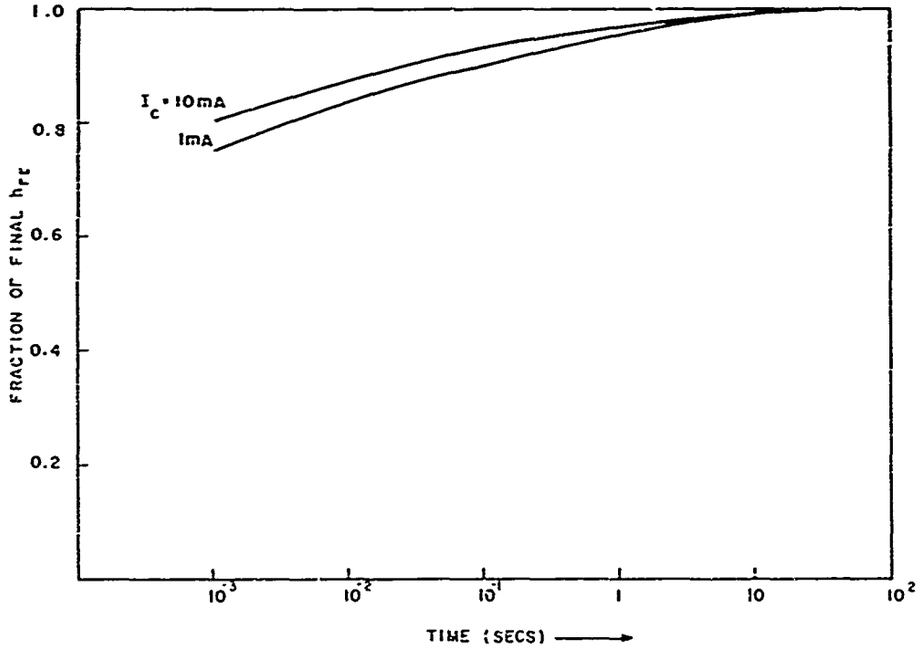


Figure 42. Recovery of A2409 After 5 x 10^{13} N/cm^2

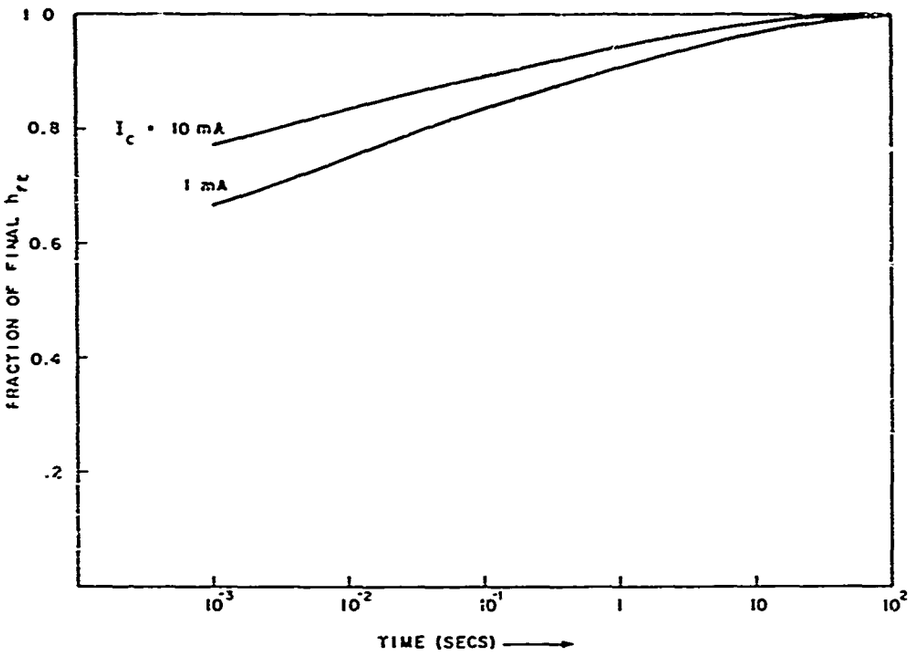


Figure 43. Recovery of MM999 After 5 x 10^{13} N/cm^2

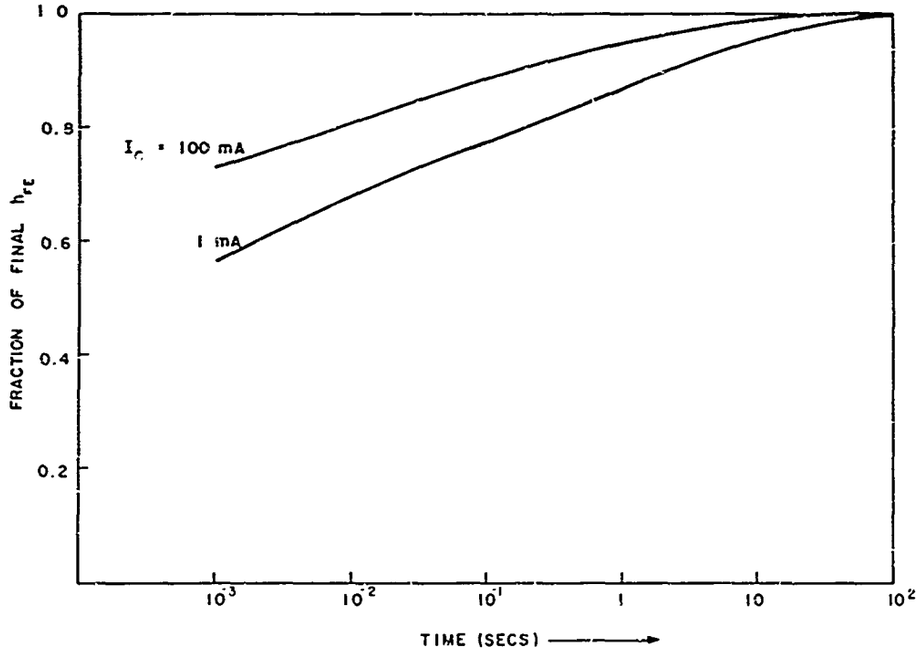


Figure 44. Recovery of 2N3252 After 5 x 10^{13} N/cm^2

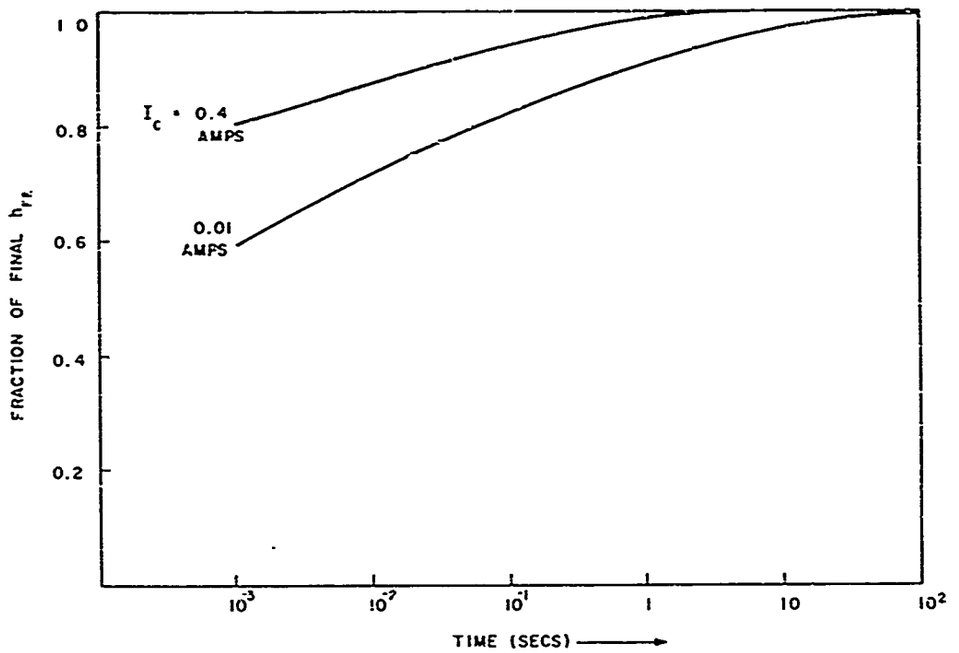


Figure 45. Recovery of SF2585A After 5 x 10^{13} N/cm^2

base region of the transistor, and it is assumed further that the centers which recover are independent of the permanent ones so that the recombination rates add, then the ratio of the recoverable recombination rate to the permanent recombination rate R_A/R_N ($R \propto 1/\beta$, where β is the d-c gain) can be determined.

$$R_A/R_N = \frac{1/\beta(t) - 1/\beta(\infty)}{1/\beta(\infty) - 1/\beta(0)} \cong \frac{I_B(t) - I_B(\infty)}{I_B(\infty) - I_B(0)} = F - 1 \qquad (1)$$

where I_B is the base current and F is the quantity defined by Sander and Gregory[1] as the annealing factor.

The quantity R_A/R_N at 1 msec has been obtained for the transistors listed above and was found to be independent of the total integrated flux of neutrons received by the device in the range from 10^{13} to 10^{14} N/cm^2. That is, the ratio of recovered to permanent damage is independent of the total fluence received. This result was also found by Sander and Gregory and supports their argument that recovery is caused by changes in centers in the bulk rather than changes in surface states.

In general, the quantity R_A / R_N for times greater than 1 msec has a time dependence of the form

$$R_A R_N = Bt^{-1/n} \tag{2}$$

where time (t) is expressed in milliseconds and the quantity $3 \leq n \leq 6$ at room temperature. Both n and B are functions of temperature and the emitter current during recovery. Thus the kinetics of recovery are described by an equation of the form

$$\frac{d(F-1)}{dt} = k(F-1)^{n-1} \tag{3}$$

and the order of recovery (n - 1) varies from 2 - 5. Nihoul and Stals[3] have discussed bimolecular recovery processes and have shown that large deviations from n=2 can generally be expected. It is therefore possible that the recovery is a result of a bimolecular diffusion process. However the strong dependence of recovery on current density has already been noted. As the current gain of the transistor changes, its current density also changes because of the emission crowding effect. Presumably this dependence could be taken into account but it will further modify the derived order (n - 1) of the recovery kinetics.

The expression equation (2) implies very large values for $R_A R_N$ for extremely short times. Unfortunately times less than 1 msec were not studied because of the long response time of the Visicorder galvonometers. Sander and Gregory's experiments[1] indicate that the maximum value of $R_A R_N$ is of the order of unity. This shows that recovery in our devices during the first msec or so is relatively small.

Sander and Gregory were able to measure F at short times by studying the recovery with very low average currents using 10 μsec pulses to measure the current gain. Their results with this technique also indicate that some recovery does take place without injection but with a very long time constant. The marked dependence of recovery on forward bias was also observed in their experiments even at very low duty cycle. Using a duty cycle of 0.1 per cent, they observed a displacement in the annealing curve to two orders of magnitude longer time compared with the result for a 100 per cent duty cycle.

From the results of Sander and Gregory and the results given in Figures 41 through 45, it is estimated that the recombination rates immediately after the burst may be as high as twice the final recovered value. That is, the neutron burst produced a recoverable recombination rate that is roughly equal to the final (fully recovered) neutron-induced recombination rate. If $1/\beta \propto \phi$, then it could also be said that the worst case degradation corresponds to the steady-state device

characteristics observed after exposure to approximately twice the actual neutron fluence. Worst case degradation should be anticipated in devices which are off during the exposure and switched on briefly some time later.

(2) Effects of Bias and Temperature on Recovery of Transistors. The SF2585A medium-power devices and MM999 devices were also operated with a collector bias of 45 volts and low collector current during irradiation; no differences in recovery phenomena were observed. All of the devices were also tested with reverse emitter bias supplied during burst. Reverse emitter bias should accentuate any radiation-induced surface degradation in the neighborhood of the emitter base junction during irradiation. Again no differences in recovery behavior were observed under these bias conditions.

A measurement was also made during reactor burst of the $V_{CE}(SAT)$ of the power transistor at I_C = 400 ma, and I_B = 60 ma. $V_{CE}(SAT)$ should be fairly sensitive to changes in the lifetime of minority carriers in the collector epitaxial region. However, no recovery of the saturation voltage was observed. The relatively high base drive used in this experiment may have obscured any changes in $V_{CE}(SAT)$ due to lifetime changes in the epitaxial region.

The high power NPN's (SF2585A) exposed with collector currents of 400 ma and the medium power NPN's (2N3252) with collector currents of 100 ma had substrate temperatures appreciably higher than 30°C. (Estimated temperatures were 80°C for the SF2585A and 60°C for the 2N3252.) The recovery of these devices with emitter current applied minutes before irradiation was compared with that in devices with emitter current applied shortly after irradiation. No significant differences in recovery were observed.

The effects of temperature on the recovery process were also examined using the hot cold mount described above. For these studies only the medium power NPN and PNP transistors were used. Recovery was observed at -50°C and -100°C and compared with room temperature results. Figure 46 shows typical results for each type of device and Table 2 summarizes the behavior of these devices in terms of the parameters given in equation (1). At low temperatures no recovery was observed until bias was applied, as was the previous finding at room temperature. However at 100°C, during the first 0.5 sec roughly 1/3 of the recoverable centers disappeared before application of forward bias. At 1 msec R_A/R_N was roughly 1/2 as large at 100°C as it was at -50°C; however, recovery with bias proceeds somewhat more slowly, if anything, at the higher temperatures. This may be because the emission is not as crowded in the emitter region at high temperatures; hence the actual current densities and the minority carrier injection levels are much lower. Because of the dependence of recovery on injection level, the rate of recovery would then be somewhat lower at this temperature.

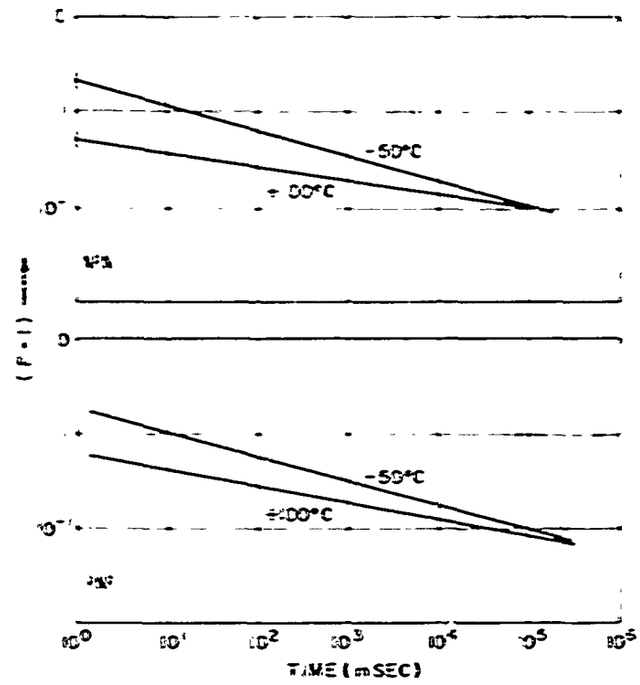


Figure 46. Effect of Temperature on Recovery of NPN and PNP Silicon Transistors

TABLE 2

Recovery Parameters of 2N3252 and 2N3244 After Exposure to Burst of Neutrons (t = o)

(R_A R_N = F - 1 = Bt^{-1} ^{n}, t given in msec)

Device	Collector Current (ma)	Temperature (°C)	B ($msec^{-1}$)	n
NPN - 2N3252	10	-50	1.2	4
	10	-25	0.9	5
	10	+100	0.6	6
PNP - 2N3244	10	-50	1.2	4
	10	-25	0.9	5
	10	+100	0.7	5

The recovery of transistors after a burst of neutrons was also monitored using an exponentially swept emitter current. The emitter current was increased from 1 μa to 20 ma in approximately 100 msec. Figure 47 shows the base current of the NPN A2409 as a function of the emitter current. It was hoped, in this type of measurement, to be able to separate space-charge region recombination currents induced by radiation from bulk recombination currents. Recovery was sufficiently rapid at the high emitter currents, so that it was not possible to sweep an entire I_B versus I_E curve before recovery took place. Because of system response, recovery for times less than 5 msec was not determined in these tests. The results for longer times give somewhat smaller annealing factors, F, than those determined using constant emitter current. This result is attributed to the relatively high average-current density used in these measurements which would speed up the recovery of the recoverable centers.

(3) Diodes. The transient recovery of forward voltage drop in a WECo G657174 PIN was also studied. The PIN diode is in theory a desirable structure for studying changes in recombination rate for several reasons: (1) its forward voltage drop is a very sensitive function of the diffusion length in the intrinsic

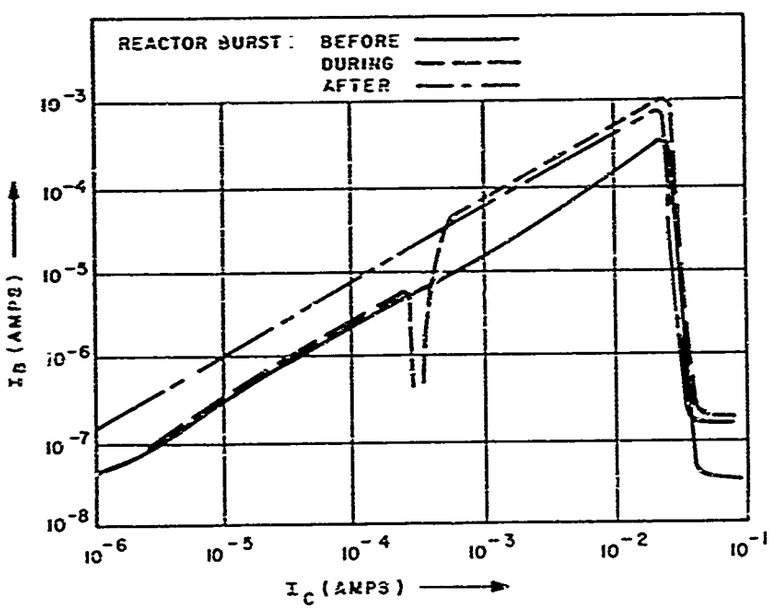


Figure 47. Recovery of Base Current (I_B) of Silicon NPN A2409 After Reactor Burst of 5×10^{13} N/cm^2.
I_C is swept from 10^{-6} to 2×10^{-2} amps in 100 msec.
Burst occurs roughly halfway through one sweep.

region, (2) its current density is uniform and calculable, and (3) its characteristics at high injection level should be relatively insensitive to surface effects. A simple theory gives the following relationship between the forward voltage drop and the diffusion length in the wide base region.

$$V = V_{PN} + V_I$$

$$= \frac{2kT}{q} \ln \left(J \sqrt{b}\, L/2q\, D_n\, n_i \right) + \frac{kT}{q} \frac{\pi}{2} e^{W/2L} \qquad (4)$$

where b is the ratio of electron and hole mobilities, L is the minority carrier diffusion length, W is the width of the intrinsic layer, D_n is the electron diffusion constant, and n_i is the intrinsic concentration. For the injection levels and ratios of width and diffusion length for which this particular device was operated, the first term in this expression gives rise to a voltage drop of the order of 0.5 - 1.0 v which depends on the current density and the temperature of the measurement. The values for the junction voltage drop assumed in this study are given in Table 3.

TABLE 3

PN Junction Voltage Assumed for WECo PIN Diode

Temperature (°C)	Current (ma)		
	1	10	50
-50	0.70	0.80	0.87
25	0.60	0.70	0.77
100	0.45	0.60	0.67

The PN junction voltage (first term in equation 4) also changes with irradiation; however, the dependence on lifetime is logarithmic and has been neglected in this analysis. Comparisons made between irradiated and unirradiated PIN's show not only the expected change in saturation current contribution to the first term, but also an unexpected change in the slope of the ln I versus V characteristic which has not been taken into account. The change in slope could result from a strong dependence of diffusion length on injection level.

The second term in equation (4) which is independent of diode current gives rise to a large additional drop when the ratio of the width of the intrinsic region of the diffusion length of minority carriers is large compared to unity. At 5×10^{13} N/cm^2 the estimated lifetime in the intrinsic layer is of the order of 0.1 μsec. Since the base width of these devices is approximately 50 μ, the ratio of base width to diffusion length is greater than three. Hence the second term should add significantly to the forward voltage drop.

Recovery of the forward voltage drop as a function of time after a single neutron burst for various diode currents is shown in Figure 48. The current density at 10 ma in this PIN is approximately 5 amp/cm^2. The recovery here is quite similiar to that of the current gain in the transistors and occurs in roughly the same time span. Essentially, complete recovery does not occur until times are of the order of 100 sec or more, possibly because of the low current density. If the PN junction voltage given in Table 3 is subtracted from these results, the voltage across the intrinsic region can be obtained. It will be seen that this voltage is 3 to 4 times as great immediately after the neutron burst as it is in a completely recovered diode. From equation (4), the voltage drop across the intrinsic layer, a simple relationship between the recoverable and the final recombination rates in the intrinsic layer similar to that obtained for the transistors can be derived.

$$\frac{R_A}{R_N} = \frac{\ln^2 \alpha \Delta V(t) - \ln^2 \alpha \Delta V(\infty)}{\ln^2 \alpha \Delta V(\infty) - \ln^2 \alpha \Delta V(o)} \qquad (5)$$

where $\alpha = q/kT$ and $\Delta V = V - V_{PN}$

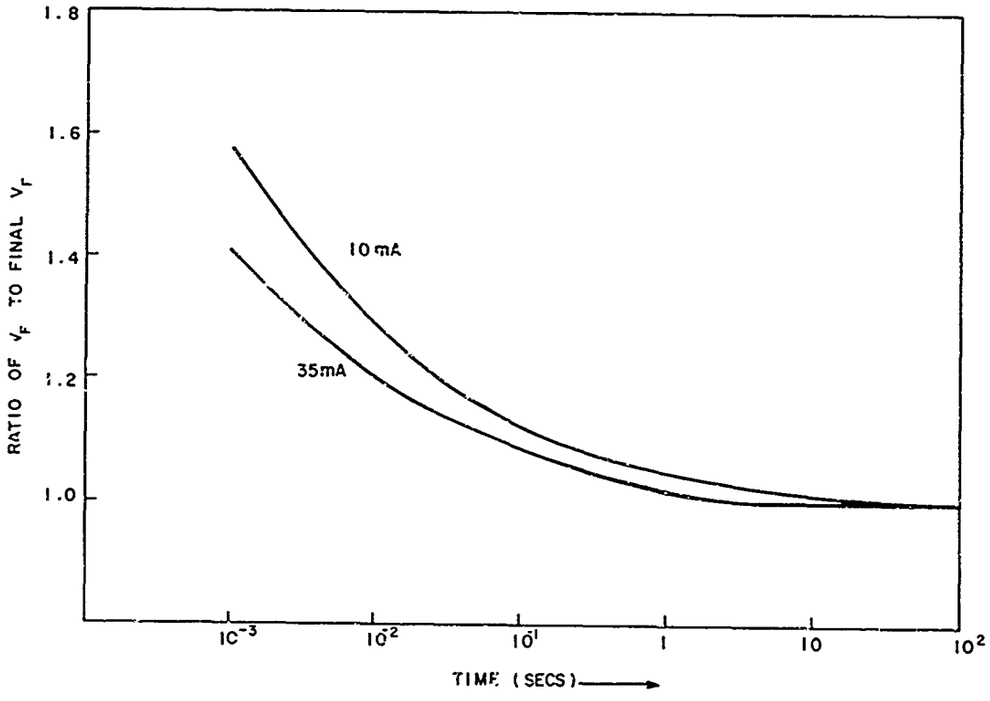


Figure 48. Recovery of Forward Voltage (V_f) of PIN Diode After 5 x 10^{13} N/cm^2

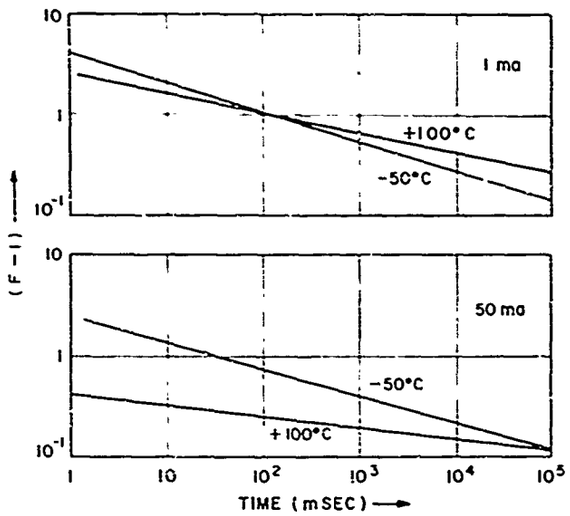


Figure 49. Effect of Temperature on Recovery of WECo PIN Diode

From this expression and the results shown in Figure 49 values of R_A/R_N at 1 msec of 1.3 at 35 ma and 1.7 at 10 ma are obtained. These results are somewhat larger than the results obtained for the transistors. The time dependence of R_A. R_N for diodes could also be fitted by an equation of the form of equation (2) with $3 \leq n \leq 6$. Again initial values of the unrecovered device voltage drops were not obtained at times less than 1 msec. Sander and Gregory have studied the recovery of solar cells at very low current densities. Their results show a flattening of recovery at times less than 1 msec. It may be concluded from their results that the max value of R_A R_N for diodes lies in the range 1 to 3.

Exposures of the PIN diode at a series of integrated flux levels also showed that in the case of diodes R_A/R_N was independent of the flux level. This result, plus the magnitude of the annealing observed in the PIN diode, is fairly strong evidence that a bulk phenomenon, not a surface-dependent effect, is being observed.

The effects of temperature on recovery in the PIN diodes were also studied and the results are shown in Figure 49. These results can again be expressed in terms of the parameters given in equation (1). Their values are summarized in Table 4. In general these results are quite similar to those obtained on transistors, i.e., the magnitude of recovery is larger at low temperatures and larger at low measuring currents. However the rate of recovery is relatively insensitive to both the temperature and measuring current density. Again, as in the case of the transistors, recovery of the PIN's was not observed until forward current was passed through the device. The results of Sander and Gregory[1] on solar cells indicate that recovery without bias (in their case, application of light to the samples) proceeds very slowly with a time constant of the order of many seconds.

TABLE 4

Recovery Parameters of WECo PIN Diode
After Exposure to Burst of Neutrons (t = o)
($R_A / R_N = F - 1 = Bt^{-1/n}$, t in msec)

Current (ma)	Temperature (°C)	B ($msec^{-1}$)	n
1	-50	3.5	4
	+25	3.0	4
	+100	2.0	5
50	-50	1.5	4
	+25	1.0	5
	+100	0.6	>6

The severe recovery effects observed in PIN diodes might also be expected in the on-voltage of PNPN devices in view of the fact that PNPN's also operate at relatively low current densities. The severity of the effect will undoubtedly be most noticeable in circuits which are off during exposure to radiation and are switched on for only brief periods after irradiation. In this case the maximum observed annealing factor must be expected.

c. Other Recovery Effects

(1) Reverse Recovery. An interesting reverse recovery effect was observed in neutron-irradiated medium-power NPN and PNP devices which were heated to anneal out some of the neutron damage. The degraded post-irradiation gain at low currents (1 ma) increased about 40 per cent after a 5-minute, 125°C anneal. This increase was slowly reduced to 25 per cent after a period of several minutes. A brief operation at high current (100 ma) would accomplish the same reverse recovery in a matter of seconds. The effect was reversible; reheating to 125°C again increased the initial gain recovery to 40 per cent. A comparable reverse recovery was observed in the forward voltage of irradiated PIN's. Thus it is unlikely that it is a surface controlled effect.

R_A/R_N has been evaluated as a function of time for reverse recovery using equation (1). Figure 50 compares the time dependence of R_A/R_N for reverse recovery of neutron-irradiated medium-power NPN after a 200°C anneal with the typical recovery of R_A/R_N of similar devices after a burst of neutrons. Note that the time dependences and magnitudes of the two effects are very similar. The reverse recovery effect like the normal recovery is largest for the lowest measuring current.

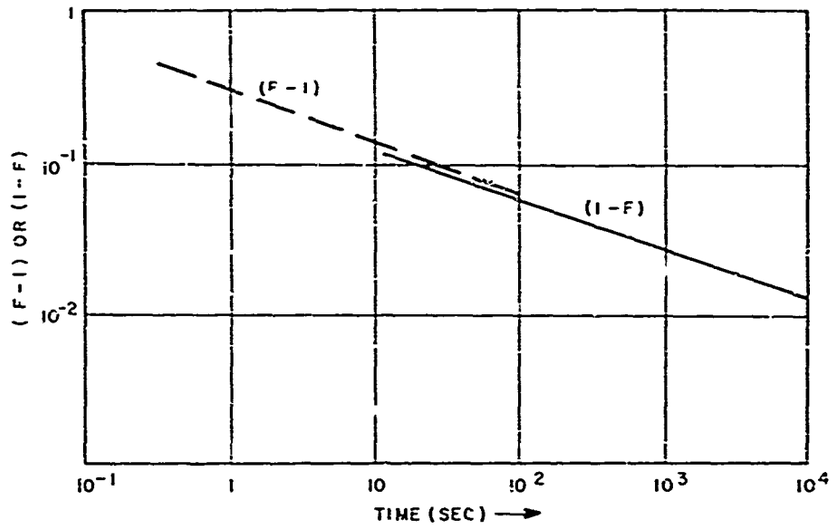


Figure 50. Comparison of Recovery of NPN 2N3252 After Burst of Neutrons (5×10^{13} N/cm^2) and Reverse Recovery (1-F) After 200°C Anneal

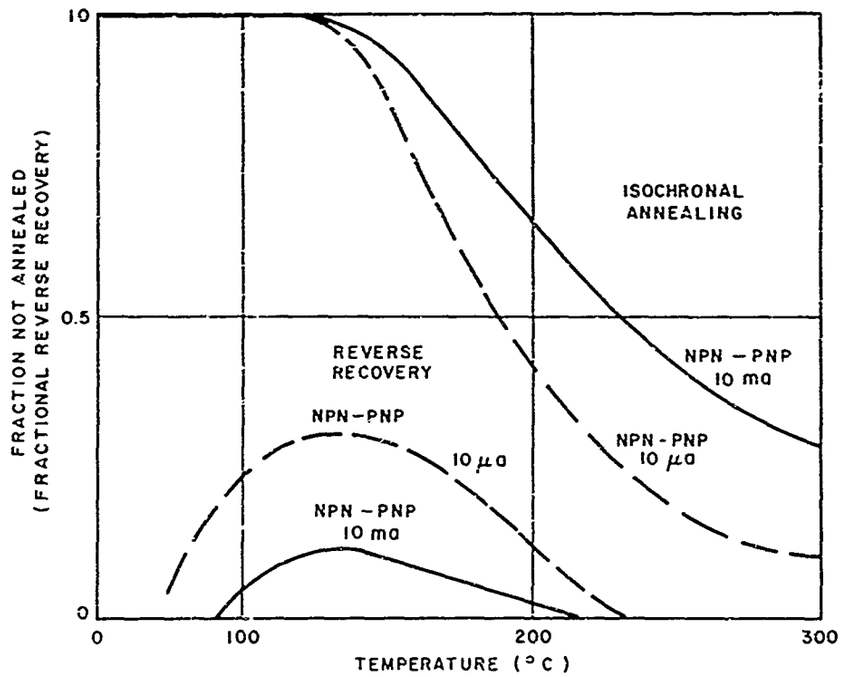


Figure 51. Isochronal Annealing and Magnitude of Reverse Recovery (1-F) versus Temperature for Silicon NPN (2N3252) and PNP (2N3244) After Fast Neutron Bombardment

Figure 51 shows how the fast neutron damage anneals for the medium-power NPN and PNP devices used in this study. Little or no anneal of damage is measured up to 150°C. The magnitude of the reverse recovery effect is also plotted in Figure 51. This effect peaks in the neighborhood of 150°C and gradually disappears in the same region, from 175° to 225°C, where most of the neutron damage anneals. It may be concluded then that the reverse recovery is associated with bulk neutron damage.

It may also be concluded that the reverse recovery is a bulk electronic process because it is reversible and because it is largest below the temperature where a significant amount of annealing of neutron damage occurs. The marked similarities of ordinary recovery, and reverse recovery in turn, suggest that ordinary recovery may also be due to an electronic process.

(2) Capacitance Recovery. Measurements of the reverse-bias junction capacitance have also been made in devices immediately after exposure to bursts of fast neutrons from the White Sands FBR. The device studied was a WECo 426G rectifier having a N^+PP^+ type of structure with a P region doping of 10^{14} cm^{-3}. Junction capacity after a neutron burst drops abruptly and slowly climbs back to nearly its pre-irradiation value. For a N^+P step junction, the capacitance variation is given by

$$C(t) = k(N_o - N_a - N_n)^{1\ 2} \tag{6}$$

where k is a constant depending on the applied bias, N_o is the initial acceptor concentration, N_a and N_n are the concentrations, respectively, of recoverable and final neutron-induced deep donors It is easily shown that

$$\frac{N_a}{N_n} = \frac{C^2_{(\infty)} - C^2_{(t)}}{C^2_{(o)} - C^2_{(\infty)}} \tag{7}$$

A-c capacitance measurements at 10 kHz and 100 kHz were made using a Boonton capacitance meter. The time dependence of N_a/N_n at a reverse bias of 0.5 volts is of the form

$$\frac{N_a}{N_n} = 4.0 \exp\left[-t/650\right] \tag{8}$$

At 500 sec, forward bias was applied and almost complete recovery of the capacitance to its final value occurred in a few seconds. Similar capacitance recovery was observed at larger values of reverse bias. The values given above were obtained with 100 kHz measuring frequency. Similar results were obtained at 10 kHz, although there are quantitative differences which are as yet unresolved.

The observed exponential dependence is characteristic of an electronic process, i.e., filling or emptying of traps. Since the defects which are responsible for the changes in a-c capacitance are located at the edge of the space charge region, they see very few mobile carriers and the charging process could be very slow. Forward injection would speed up the charging process drastically.

From the final capacitance change it is estimated that N_n is of the order of $10^{13}/cm^3$ for a fast neutron-fluence of 5×10^{13} N cm^2. This corresponds to roughly one or two acceptors being neutralized per cluster. The initial value $N_n + N_a$ immediately after the burst is five times as large, hence a large number of charge states appear to be associated with each cluster.

There is as yet no evidence to link capacitance recovery and the post neutron burst recovery of transistor gain, except the common dependence of recovery rate on injection level. However, there is every reason to believe that changes in the charge state of a cluster could alter both its effectiveness as a recombination center and the effectiveness of other recombination centers located within a Debye length of the cluster.

d. Discussion

(1) Nature of Neutron Damage in Silicon. Unfortunately there is little solid experimental knowledge about the radiation damage centers produced by fast neutrons in silicon. As a result, the model discussed below for neutron damage is a highly speculative one.

For fission spectrum neutrons the average neutron energy is of the order of 1 Mev. The mean free path of such a neutron is approximately 6 cm in silicon; hence only 18 per cent of the incident neutron flux produces recoils in a 1-cm-thick sample of material. The average energy transferred by this neutron to a primary silicon recoil is of the order of 70 kev. In the silicon lattice this recoil can travel a linear distance of the order of 600 Å and give up 70 per cent of its energy to subsequent displacements. If the average energy per displacement is assumed to be 25 ev, then the total number of displacements will be $\approx$ 2000. Of these displacements a large fraction can be assumed to be close interstitial-vacancy pairs (Fraenkel defects) which recombine almost immediately. Of the remaining defects it is generally assumed that the interstitials (transferred by replacement collisions) exist as a cloud surrounding a core with high vacancy concentration. The overall defect is commonly called a cluster.

The cloud of interstitials might be expected to rapidly collapse because of the large interstitial mobility, but trapping by other lattice defects or by other interstitials could slow down considerably the return of interstitials to the vacancy core.

Within the core, rapid rearrangement of the vacancies to form vacancy pairs, triplets, and possibly vacancy platelets can be expected. Interactions of the vacancies with other defects within the core (oxygen donor or acceptors) should be less probable because their concentrations are comparatively small.

The result after the first immediate stage of annealing should consist of a core, perhaps 300 Å in diameter, which has several hundred vacancies aggregated into vacancy pairs, triplets, etc., surrounded by trapped interstitials and possibly a small density of trapped vacancies (bound to oxygen, donors, etc.). This model is illustrated in Figure 52.

Several experiments in germanium, including electron microscopy,[4,5] seemed to support this type of model for neutron damage but as yet no evidence exists for this model in silicon.[6]

It may be concluded from this kind of model that two possible types of damage could be measured in fast neutron-irradiated materials. The first is diffuse damage arising from the cloud of impurity-bound vacancies and interstitials surrounding the core. This diffuse damage should be very similar to the damage produced by electrons or photons having energies less than a few Mev. Electrical measurements of electron and photon-bombarded silicon indicate that isolated vacancy-oxygen,

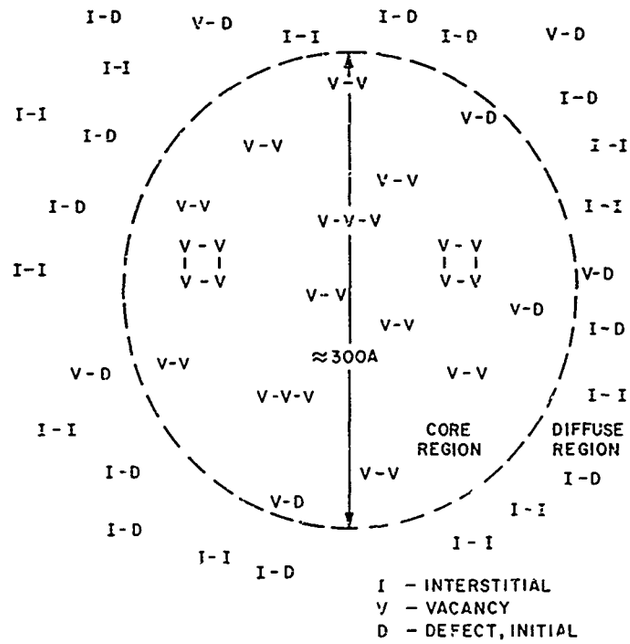


Figure 52. Cluster Model for Fast Neutron Damage

vacancy-donor, and vacancy-vacancy complexes are the most significant contributors to this type of damage. The second kind of damage is due to the vacancy rich core of the cluster. This kind of damage would not be produced in low energy electron or photon irradiations.

(2) Electrical Effects of Neutron Damage in Silicon. Past neutron damage in silicon has been studied using EPR, optical, Hall, and lifetime measurements. Only the EPR studies can be interpreted in terms of specific kinds of defect centers. The experiments of Jung and Newell[7] have been interpreted by Corbett.[8] The experiments indicate relatively large concentrations of vacancy aggregates, including vacancy pairs and quadruplets. Vacancy-oxygen and vacancy-donor pairs have not been observed.

Optical measurements indicate divacancy and some vacancy-oxygen complexes. Lifetime, Hall, and photo-conductivity measurements indicate defect levels at 160 mv below the conduction band and 270 mv above the valence band. These levels have been identified respectively with the vacancy-oxygen and the divacancy complexes. In addition, many other levels whose positive identification is highly uncertain have been observed. R. Whan[9] has observed that the vacancy-oxygen optical-absorption band begins to develop slowly in neutron-irradiated material at temperatures in excess of 50 C. This slow increase in absorption is attributed to the breakup of the vacancy-rich core and subsequent migration of the freed vacancies to oxygen traps.[9]

Two authors[10,11] have observed that neutron-irradiated silicon exhibits lifetime degradation which is relatively independent of the impurity content of the material in contrast to very marked impurity dependence of lifetime degradation in gamma- or electron-bombarded materials. This suggests that the principal recombination centers are the vacancy cores of the neutron clusters.

H. Stein[12] has observed a very small temperature dependence of the production rate of neutron lifetime damage in bulk material. (This observation has been confirmed in a variety of silicon devices as well.) Again, this result contrasts with the lifetime production rate observed in electron- or gamma-radiated materials where there is a marked temperature dependence. Stein also observed recovery of the lifetime in neutron-irradiated materials at temperatures 50 C which he also attributes to breakup of the vacancy cores.

Annealing of neutron damage is typically spread out over a broad range of temperatures, as shown in Figure 52. On the other hand vacancy-oxygen, vacancy-donor, and vacancy-vacancy complexes have been observed to anneal rather sharply in a very narrow temperature range in electron- or photon-irradiated materials. For example, the vacancy-phosphorus center anneals almost completely between

150°C - 190°C, whereas the vacancy-oxygen and vacancy-vacancy centers anneal rather abruptly in the range of 300°C. The shape of the neutron damage annealing curve has been attributed to the breakup of the vacancy-rich clusters over a wide range of temperatures.

The general conclusion to be drawn from these various observations is that the effects of neutron damage on semiconductor electrical properties are quite different from the effects of the diffuse type of damage observed in electron or gamma environments. The specific conclusion that some observers have drawn[10,12] is that the electrical changes in neutron-bombarded silicon are primarily related to the vacancy-rich core of the cluster.

(3) Cluster Electronic Model. B. R. Gossick[13] has proposed a model for the electronic effects of the disordered regions in semiconductors resulting from fast neutron bombardment. Within the disordered regions in silicon, the material is presumably near intrinsic since the disorder introduces states near the middle of the forbidden gap. To some extent the neutron disordered regions resemble disordered regions present around dislocations. V. Heine[14] has argued that the new states introduced into the forbidden gap should be acceptors, i.e., negatively charged when they capture an electron. These states are removed from the conduction band because of the dilation of the lattice within the disordered regions. Gossick, however, assumes that both donor and acceptor states are introduced, so that in either P- or N-type material an island of nearly intrinsic material can be formed. The divacancy, which should be present in large concentrations in the core, has been shown by Watkins and Corbett[15] to have some of the necessary properties for this purpose, i.e., the divacancy can bind either a hole or an electron with a fairly deep binding energy. Since the Fermi level position in the disordered region immediately after it is created is different from the surrounding material, a potential must develop between this region and the surrounding material. The potential is developed across a dipole layer which builds up at the edge of the disordered region. The dipole layer is formed by trapped holes or electrons within the disordered region and by ionized acceptors or donors outside the disordered region. The dipole space charge region extends into the bulk material a distance of the order of a Debye-length (L)

$$L = 1/q\ (kT\epsilon/N)^{1/2} \tag{9}$$

where ϵ is the dielectric constant and N is the impurity concentration in the bulk. L is the order of 500 Å for 1.0- and 5.0-ohm-cm N- and P-type silicon, respectively.

The potential drop required between the disordered and bulk regions to bring their Fermi levels into equilibrium will be about a few tenths of a volt. Assuming the cluster radius is 150 A, the capacitance (C) of the cluster is estimated to be

$$C = 4\pi\epsilon \ (150 \times 10^{-10})$$

$$\simeq 1.7 \times 10^{-17} \text{ farads}$$

The number of charges (n) required to produce a potential drop of 0.2 v would then be

$$n = cv\ q$$

$$n \simeq 20$$

Thus 20 electrons or holes must be captured by the disordered region to bring it into equilibrium with the surrounding material. This model is illustrated in Figure 53. In heavily doped material the equilibrium situation should be achieved rapidly. In lightly doped material (or junction depletion regions) equilibrium between the disordered region and the bulk will be reached very slowly since

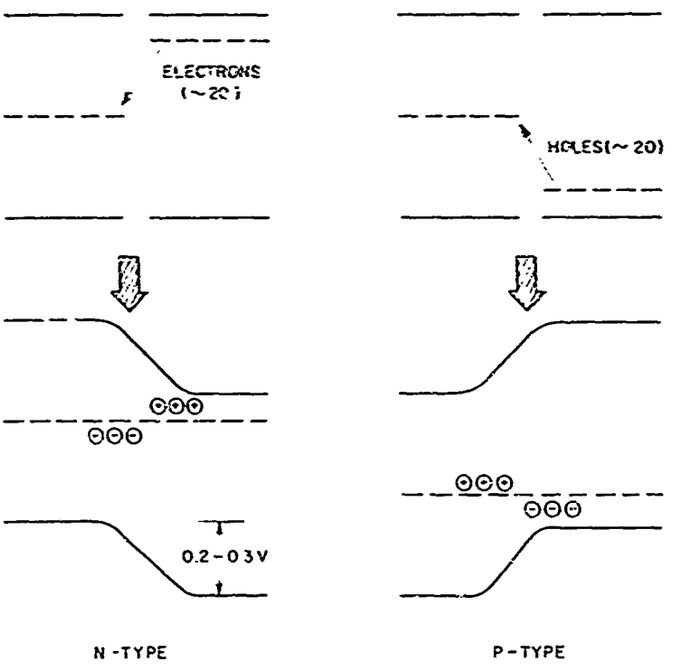


Figure 53. Gossick Model for Electronic Properties of Neutron Clusters

only a small fraction of the majority carriers encountering the disordered regior. will be energetic enough to overcome the coulombic repulsion of the center; i.e., the capture cross-section of the cluster for majority carriers becomes very small. On the other hand minority carriers will see an attractive potential near the cluster and their capture cross-sections ought to be very large. In other words the clusters are primarily minority carrier traps. The following experimental observation supports this conclusion. At 5×10^{13} N/cm^2 lifetimes typically measured in silicon devices are of the order of 0.1 - 0.2 μsecs. If the capture cross-sections for both majority and minority carriers are assumed to be of the order of the geometrical size of the cluster, i.e., 2×10^{-11} cm^2, then the bulk lifetime can be estimated at 5×10^{13} N/cm^2 to be

$$\tau = \frac{1}{N\sigma v} \cong 0.001 \text{ μsecs}$$

where v is the thermal velocity of the minority carriers. Thus at least one of the capture cross-sections must be much smaller than the geometric cross-section. Since it is physically unlikely that both cross-sections are small, it is concluded that clusters, particularly large clusters, should be traps. A large number of experimental observations[15,16] in neutron-irradiated silicon suggest that large trap concentrations are present. However no attempt has, as yet, been made to relate such traps to the clusters themselves. Wysocki[17] has observed thermally stimulated conductivity from traps roughly in the middle of the gap. Strong evidence of traps is also seen in plastically deformed silicon.[18]

If it is concluded that large clusters are primarily traps, then the increased recombination from fast neutron bombardment must be accounted for. The following possibilities are suggested.

1. Small (weakly charged) clusters.
2. Diffuse defects surrounding cluster.
3. Recombination-generation in space-charge region of cluster.

Since the results quoted previously show that neutron damage in silicon does not resemble gamma or electron irradiation damage, the second possibility must be rejected. Similarity, the third possibility would suggest a strong dependence of lifetime damage on the material doping. This dependence has not been observed.

It is concluded then that small clustered regions which have relatively few holes or electrons to attain equilibrium may account for increased recombination-generation associated with neutron damage but large centers, presumably due to

the more energetic collisions, will behave as traps. The ratio of traps to recombination centers would be expected to increase as the neutron energy increased.*

(4) Models for Transient Recovery. In the analysis of the transient recovery of both diodes and transistors the recovery was expressed in terms of R_A/R_N. This ratio can be related to the defects responsible for recovery in the following way,

$$R_A/R_N = \frac{N_A \sigma_A}{N_N \sigma_N} \tag{16}$$

where N_A, N_N are the numbers of recoverable and permanent centers respectively and σ_A and σ_N are their respective limiting capture cross-sections. The time dependence of R_A may be due to a change in concentration of centers, in their capture cross-section, or both. The number of recombination centers can be changed by atomic migration of defects (annealing), whereas the cross-section can be altered by changes in their charge state, i.e., by electronic processes.

Sander and Gregory[1] and Binder and Butcher[2] have interpreted the ordinary recovery process in terms of atomic migration, specifically the migration of interstitials back to the vacancy-rich cluster. The vacancy complexes within the core are presumed then to break up, releasing some vacancies outside the cluster. It is the release of these vacancies that is believed to give rise to the slow growth of vacancy-oxygen centers observed by Whan.[5]

There are a large number of experimental observations which do not support this explanation for the recovery process. Among these are the following.

1. In devices, little or no annealing of lifetime degradation is observed in the temperature range from -50°C to +150°C (see Figure 51).
2. Such an interpretation fails to account for the marked dependence of the recovery process on minority carrier injection.†
3. In the case of atomic migration a considerable change in the rate of recovery with temperature would be expected. This was not observed.

---

*The Gossick Model roughly explains the carrier removal rates of fast neutrons in either n or p type silicon. Thus for 1 N/cm^2 one expects 0.18 clusters/cm^3 and 20 x 0.18, i.e., 3.6 carriers/cm^3 removed per neutron. Measured carrier removal rates are of this order and are approximately equal in n and p type material.

†Sander and Gregory suggest that interstitial mobility (like vacancy mobility) may depend on charge state, hence on injection level. If this were the cause of the rapid recovery under forward bias it would be very difficult to explain the similarity of results obtained on NPN and PNP transistors.

4. This interpretation could not account for the reverse recovery process discussed in Section 4. This process has many features in common with the ordinary recovery process.

As a result of these experimental observations, an alternative model for the recovery process, based on tentative conclusions given in the previous section, is proposed.

These conclusions are:

1. Large multi-charged clusters are minority carrier traps.
2. Small and weakly charged clusters are recombination centers.
3. Charge equilibrium for large clusters is attained slowly in lightly doped material or in junction space-charge regions.

It is expected that following a burst of fast neutrons both the small and large clusters would be weakly charged in lightly doped (I) material or in junction space-charge regions. Hence, both small and large clusters could be recombination centers. As the large clusters gradually charge-up, their cross sections for majority carrier capture decrease and these clusters slowly convert to traps.* It is assumed that within the cluster there are both the deep trap levels called for in the Gossick model, and deep or shallow recombination levels. Thus the recovery process is considered to be associated with slow transformation of large clusters from effective recombination regions to traps.† An alternative way of presenting the model is to regard the edge of the cluster as a surface with both fast and slow states, where the charge in the slow states controls the recombination through the fast states.

In a typical transistor the only regions where this electronic process could be long-lived would be the emitter and collector space-charge regions. In the emitter space-charge region it would give rise to large initial space-charge recombination current, i.e., to an initially low value of emitter efficiency. This would explain the dependence of the magnitude of recovery on measuring current

*One can expect to see such a process only with fast neutron bombardment. Energetic charged particle bombardments would produce such intense local ionization that any clustered regions produced should reach charge equilibrium rapidly.

†The slow charge equilibrium of large clusters could affect the recombination rate in several other ways as well. For example, assume the diffuse defects around a cluster were primarily responsible for recombination; as the cluster charged up the band edges within a Debye length of the cluster would be bent and the recombination rate of centers inside this Debye length would slowly change.

since emitter efficiency tends to limit the current gain at low currents, whereas base transport efficiency limits current gain at high currents. Obviously recovery would be highly accelerated if the majority carrier concentration were raised in the space-charge region as it would be by a high level of minority carrier injection.

In the PIN diode the lifetime throughout the entire intrinsic region would recover to its final value slowly, as was observed.

Raising the temperature of either the transistors or PIN diodes would not markedly accelerate the recovery process because in either case the majority carrier concentration in the depletion region or the I region (not truly intrinsic, possibly 10^{12} majority carriers present) is not much altered by the temperature change.

Reverse recovery in transistors and diodes is also explained by this model in the following way. After a period of time at elevated temperature, a larger number of charges is trapped by the cluster than is required for room temperature equilibrium, i.e., more carriers have enough thermal energy to surmount the cluster potential. On cooling and injecting minority carriers, the clusters are observed to gradually lose rather than gain charge. Hence, the recombination rate slowly goes up (reverse recovery) rather than down. This process is expected to be reversible.

Although the capacitance recovery has yet to be related to ordinary recovery, both its time dependence (exponential) and magnitude (4-5 charges cluster) suggest that the process is electronic rather than atomic in nature and may also be related to changes in charge density of clusters.

e. Conclusions

Experimental results of transient recovery in high-frequency silicon devices are in substantial agreement with those reported by Sander and Gregory.[1] In general the results can be summarized by saying the worse case device degradation at a given neutron fluence (Φ) corresponds to device parameters measured at long times ($T > 10^3$ sec) after an exposure to 2Φ. (This result is based on the assumption that both R_A and R_n are proportional to Φ.)

The worse case degradation will be most evident in devices which are typically marginal in a neutron environment, high power transistors, PNPN devices, PIN diodes, etc. Largest recovery effects are observed in components having low injection levels and short duty cycle applications. PNP and NPN silicon devices

show comparable recovery effects. Low temperatures accentuate and high temperature decrease the magnitude of the recovery.

Reverse recovery of transistor gain following high temperature anneals has been observed. This process is reversible and is believed to be due to changes in the charge states of deep traps. Recovery of junction capacitance has also been observed following exposure to a pulse of neutrons. As yet this capacitance recovery has not been related to the ordinary recovery process.

Two models have been discussed to explain the observed recovery effects. The first model, which was proposed by Sander and Gregory, attributes recovery to migration of defects around neutron-induced clusters back to the core of the cluster. The second model, which is proposed in this report, attributes recovery to conversion of clusters from recombination centers to traps by slow charging of the center. The second model accounts for the dependence of the recovery amplitude on injection level, the virtual absence of recovery in the absence of injection, and for the reverse recovery effect. The first model fails to account for any of these features.

Progress and understanding the problem of transient recovery will most likely come from a study of recovery in bulk materials and from experiments which disclose a more concrete model for neutron damage.

11. HIGH INJECTION LEVEL EFFECTS IN IRRADIATED JUNCTION DEVICES

The degradation of current gain in a transistor because of increased recombination in the neutral base and base-emitter space-charge regions is the principal effect of radiation damage at low and medium currents.

At high currents other effects caused by radiation damage centers are observed. Figure 54 shows how these effects arise. In a transistor the base current flows perpendicular to the collector current. Because of the narrow basewidths used in these devices, the base current encounters considerable resistance. A lateral voltage drop which reduces the forward bias at the center of each emitter stripe and makes only the periphery of the emitter effective in carrying current develops underneath the emitter stripes. As a consequence, emitter and collector current densities can be quite high.

A simple calculation for the emitter current density J is shown in Figure 54. J increases with emitter current I_e, with base current $(1-\alpha)\, I_e$, and with base resistivity ρ, and it decreases with increasing basewidth W and emitter perimeter h.

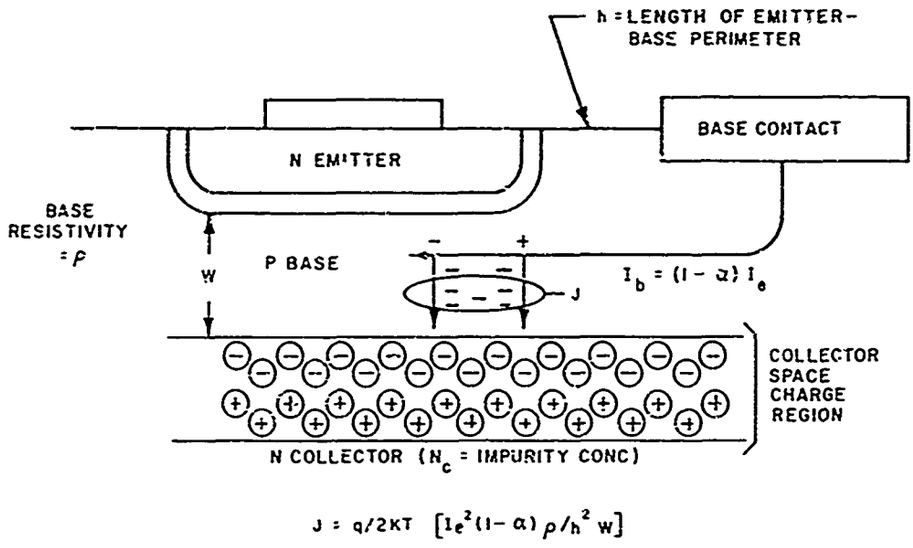


Figure 54. Emission Crowding and High-Injection Level Effects in Transistors

For the low power devices where W = 1.., h = 25i.., ρ = 1 ohm-cm and α = 0.99, an emitter current of 10 ma results in a pre-irradiation current density of approximately 20 amp/cm^2. After exposure to 2 x 10^{14}N/cm^2, base current increases by more than an order of magnitude, due to the reduction in alpha, and the emitter current density will approach 1000 amp/cm^2. In medium power devices with approximately the same constants except that emitter perimeter is about 0.1 inch, an emitter current of 1 amp after radiation again results in a current density of 1000 amp cm^2. At such large current densities various high-injection level effects set in. The crowding of the emitter current into a very small region beneath the emitter perimeter results in a smaller area for the generation of heat. This gives rise in higher thermal impedance and collector temperatures due to poorer coupling to the heat sink, and as a result, the secondary breakdown problem can be expected to be more severe.

The major effect of the increased current density is a widening of the base-width due to the "Kirk Effect,"[19] which is also illustrated in Figure 54. In this figure the mobile minority carriers have been represented with minus signs, and fixed charges in the collector space charge region have been shown by circled charge signs. The "Kirk Effect" occurs when the injected minority charge density

in the vicinity of the collector space charge region becomes comparable with the fixed charge density found there, i.e., when $J_K = qV_{max}N_c$ where V_{max} is the maximum velocity imposed on minority carriers in the space charge region by scattering, and N_c indicates the level of collector doping. This occurs in typical silicon devices at current densities of the order of 1000 amp cm^2. The heavy concentration of minority carriers tends to increase the effective charge density in the space charge region next to the base and to decrease it next to the collector. This causes the space charge region to shift toward the collector body and thus to widen the base. The wider base results in a lower value for the frequency cutoff, f_t, and in reduced dc current gain. It is expected that the effect would occur at smaller currents in more lightly doped material due to the smaller density of fixed charges. Also the effect would be expected to be worse after irradiation because the emitter crowding is enhanced by the increased base current.

From these results it can be seen that high gain at high current and at high integrated flux levels requires heavy collector doping. The optimum transistor design must, of course, strike a balance between all the factors indicated in Figure 54, such as emitter perimeter, base width and resistivity, collector doping, etc. However, it appears that a high-current, radiation-resistant transistor of conventional design is necessarily a rather low voltage device.

Generally high current operation in transistors takes place under saturation conditions, i.e., with the collector junction driven to forward bias. The high injection level effects in a saturated transistor are not easily analyzed because an appreciable fraction of the current which flows out of the base contact flows laterally in the collector region. This current gives rise to a transverse collector voltage drop similar to the transverse base voltage drop. This problem and the effects of radiation damage on the saturation operation of silicon-planar epitaxial transistors are discussed in Attachment II by E. A. Overstreet.

The conclusions reached in this study are similar to those arrived at above for normal operation of a transistor. The $V_{CE(sat)}$ which often controls the radiation resistance of an irradiated power transistor is minimized by:

1. Increasing the impurity content in the epitaxial region and reducing the width of the epitaxial layer.
2. Decreasing the sheet resistance of the base diffusion.
3. Increasing the emitter perimeter.

The conclusion is again reached that high current transistors intended for either normal or saturated operation after irradiation (because of the heavy doping required in both base and collector regions) are necessarily low voltage devices.

A further conclusion of this study is that true saturation may not be achieved in irradiated transistors unless the base current is increased well beyond the point where the collector junction voltage reverses. On the other hand, before irradiation very little is gained by increasing the base current much above this level.

There is, however, a possible way of obtaining both high voltage and high current capability in a semiconductor device which must withstand radiation, and that is to use a PNPN transistor. The principal advantage of this device is that the base current flows parallel rather than perpendicular to collector current. Consequently there is no emission crowding, and the full area of the device remains active. Another advantage is that in the PNPN device the carriers are accelerated by electric fields in their passage through the base regions instead of depending entirely on diffusion. Hence, thicker bases can be used for the same radiation resistance and the problem of high breakdown voltage, which requires wide bases, is eased. In Attachment III D.K. Wilson and H.S. Lee discuss the effects of neutron bombardment on narrow base PNPN devices.

The conclusion of their study is that such devices are capable of switching more than an order of magnitude higher volt-ampere product than transistors for the same radiation exposure. Existing devices capable of switching more than 1000 watts at 10^{15} N/cm^2 are described, as well as a proposed device structure capable of switching over 100 watts at 10^{16} N/cm^2.

12. LOGARITHMIC CURVE TRACER

a. Description

In the study of radiation damage mechanisms in diodes and transistors, a knowledge of voltage-current characteristics over a wide range of currents is needed. Important clues can be missed if the data are obtained only at discrete points rather than as a continuous curve. Since the currents in diodes and transistors are exponential functions of the diode and emitter-base voltages respectively, any wide-range presentation is best made using a logarithmic current scale and a linear voltage scale. Plots of collector and base currents in transistors should be made simultaneously to avoid errors which can be caused by temperature changes between measurements or by possible changes in characteristics after one of the currents has been swept over a wide range.

The slope of such curves as log I_C and log I_B vs V_{EB} has a value of q/nKT; if this slope is readily determinable, much useful information can be extracted from the curves. For example, one can separate the various components that contribute to the base current.

In the subsequent discussion, a logarithmic curve tracer will be described. When attached to a two-pen X-Y recorder as shown in Figure 55, it can automatically plot, in a time of the order of 10 seconds, the logarithm of one or two currents over a range of 10^{-8} to 10^{-1} ampere vs a linear voltage. At a somewhat slower tracing rate, currents down to 10^{-9} ampere may be plotted; and at a very slow rate, it is possible to plot currents down to about 10^{-10} ampere.

To date, the curve tracer has been used in surface radiation damage studies on Schottky barrier diodes, in annealing studies of neutron bombarded transistors, and in the selection of diodes and transistors to obtain devices having a wide range constant slope log I vs V characteristic.

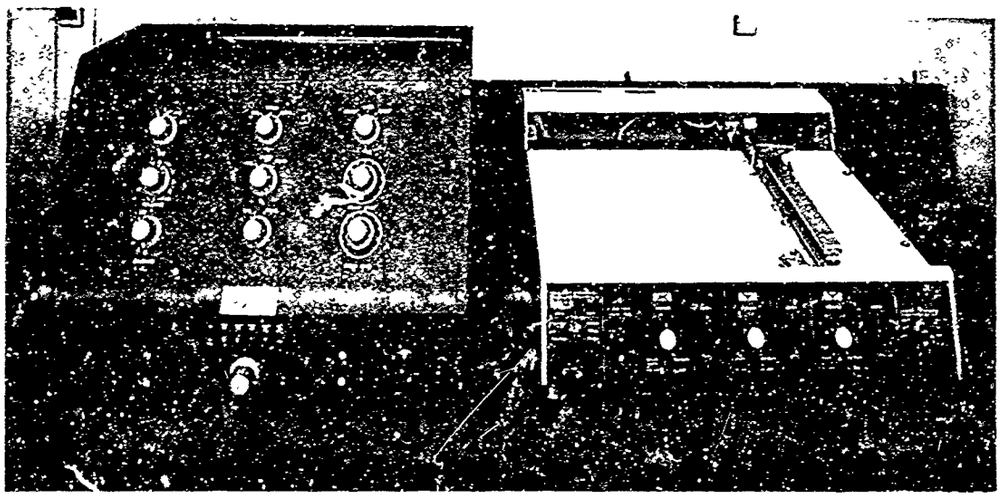


Figure 55. Logarithmic Curve Tracer

b. Basic Circuit

The basic circuit of the curve tracer is shown in Figure 56. The device under test, a transistor, is operated with its base at ground potential, its collector at a chosen V_{CB} of a few volts, and then its emitter-base voltage, V_{EB}, is varied. The resulting collector and base currents flow into separate log converters which generate output voltages proportional to the logarithms of their input currents. If these voltages are applied to an X-Y recorder, plots of log I_C and log I_B vs V_{EB} may be made as shown on the left-hand recorder in Figure 56. If the voltage representing log I_B is subtracted from that representing log I_C, a curve may be plotted representing grounded emitter current gain, I_C/I_B, vs V_{EB} as shown at the right of Figure 56.

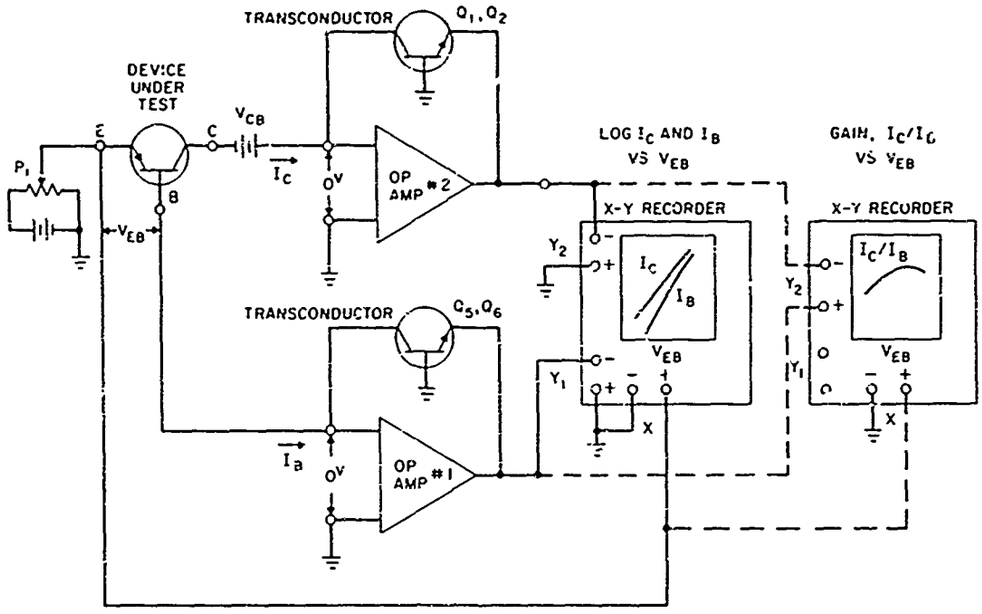


Figure 56. Logarithmic Curve Tracer Using Operational Amplifiers

c. Log Converter

In order to obtain a voltage, V, proportional to log I, a transconductor is connected between the output and inverting input of a high-gain operational amplifier as shown in Figure 56. The amplifier is chosen for low offset voltage, low offset current, and low drift. The transconductor may be either a diode or a transistor connected as a diode (base and collector common) or a transistor connected as shown. It must have a voltage associated with it which is proportional to log I over a wide current range - i.e., over as many decades as possible.

Since a high-gain amplifier with parallel negative feedback tends to maintain zero volts across its input terminals, any current fed to the input must be carried from the input to the output terminals via the transconductor. Thus, if the input current is to be 100 ma, the amplifier output must have an output capability of 100 ma. In the Philbrick line of solid-state operational amplifiers, the SP2A Differential amplifier has the desired input characteristics but an output current of only 2 ma. Hence, it is followed by a non-inverting P66A Booster amplifier as shown in Figure 57 to obtain an output current of 100 ma. The resistors and capacitors associated with the input, output, and feedback circuits are used to prevent oscillation.

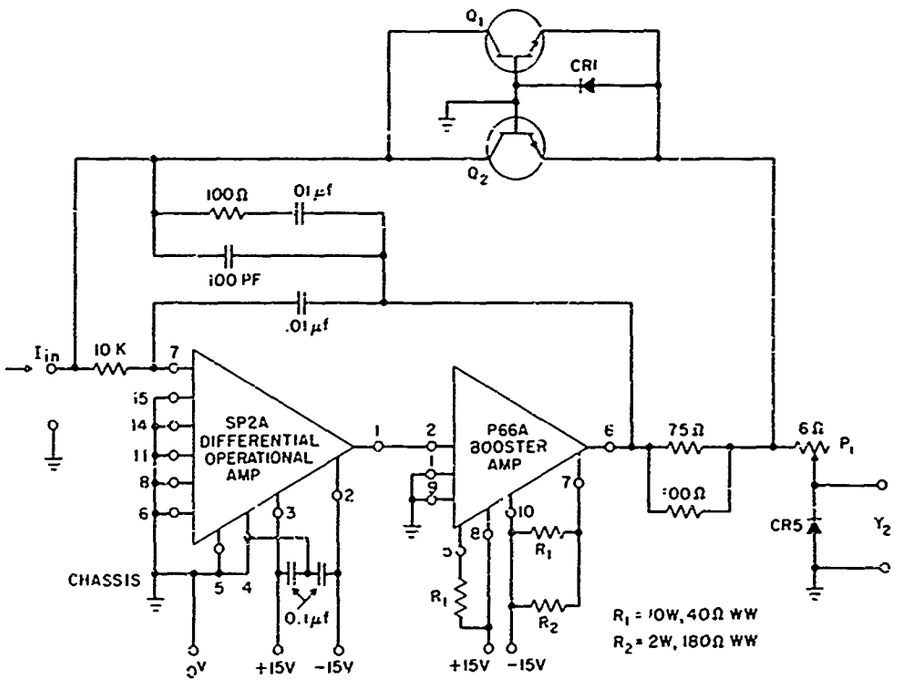


Figure 57. Details of Log Conversion Circuit

Many types of diodes and transistors (and the latter with all possible connections) were tried for the transconductor. The Philbrick PL1N and PL1P units sold for this purpose are good only to one ma. No single diode type was found which followed the E = K log I law over a range of 10^{-10} to 10^{-1} ampere. At least two diodes would have been required to cover this range.

Transistors showed more promise as transconductors and the best types tested were Motorola 2N3252 (NPN) and 2N3244 or 2N3245 (PNP) medium power transistors. By paralleling two of these as shown in Figure 57, a uniform log conversion was obtained from about 10^{-10} to about 2×10^{-2} amperes. At higher currents up to 100 ma, the voltage output of the converter increased too rapidly to maintain a straight line of I vs E on semilog paper (see Figure 58). This increase in V_{EB} above the logarithmic dependence on I is due to the flow of base current through the lateral base resistance. However, by making use of the forward V-I characteristic of a type 1N647 silicon diode, CR5, and a small variable series resistor, P_1, the curve between 10 and 100 ma may be straightened without appreciable effects at lower currents.

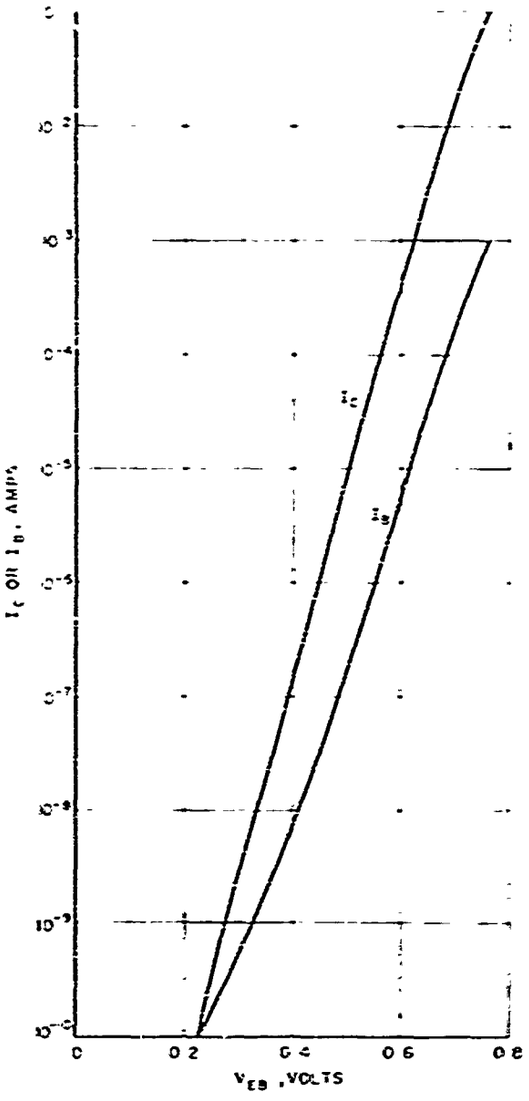


Figure 58. I_C or I_B versus V_{EB}

For best results, transistors Q_1 and Q_2 should be provided with heat sinks, but extremely good electrical insulation from cans and collector leads to ground is also required. The base-emitter junctions of Q_1 and Q_2 are protected by a type 1N647 silicon diode, CR1, from breakdown in case the wrong polarity of input current is connected to the log converter.

d. Linear V_{EB} Generator

The production of smooth I_C and I_B curves on the X-Y recorder requires stepless control of the increasing voltage, V_{EB}, applied to the device under test. This is readily attained electronically using operational amplifiers as shown in Figure 59. The circuit uses a P65A voltage amplifier followed by a non-inverting P66A current booster. A 0.001 Mf capacitor from output to input suppresses oscillation. An adjustable constant current through P_5 flows to the input of the P65A. The output potential of the P66A amplifier changes sufficiently to cause nearly all of this current to flow through the 100-uf capacitor, C_2, to the output of the P66A. Since C_2 charges linearly with time, a linear voltage ramp is generated at the output of the booster amplifier from which 100 ma can be drawn for the device under test.

With switch S6 in the Start position, P5 is set for a suitable sweep rate and P6 for the lowest value of V_{EB} to be plotted. Rotating S6 to the Sweep position causes V_{EB} to increase at the selected rate. At any value of V_{EB} along the curve the sweep can be stopped and held by using the hold position on S6. When switch

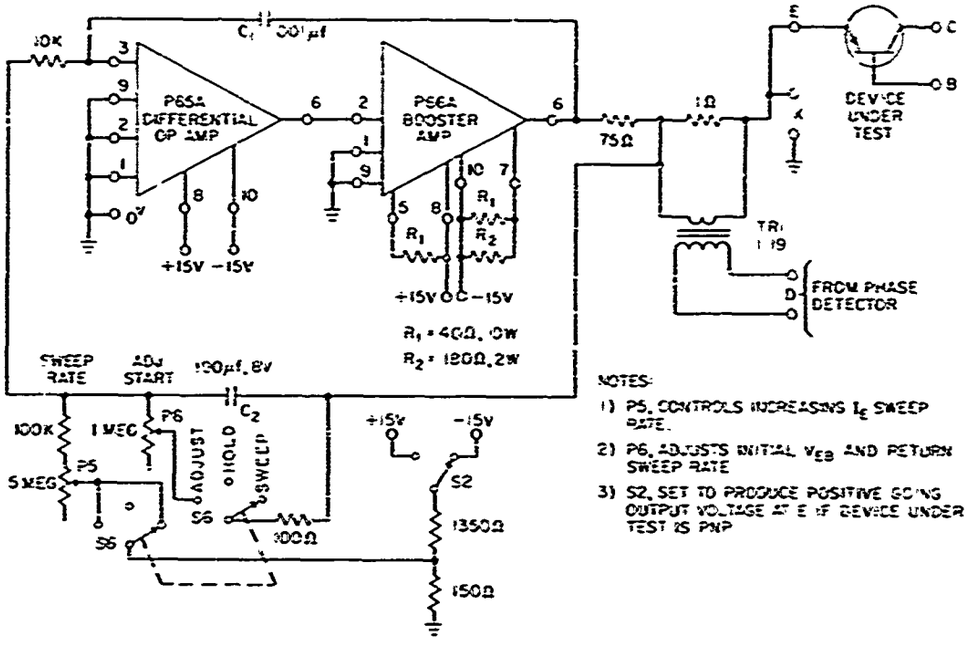


Figure 59. Linear Electronic V_{EB} Generator

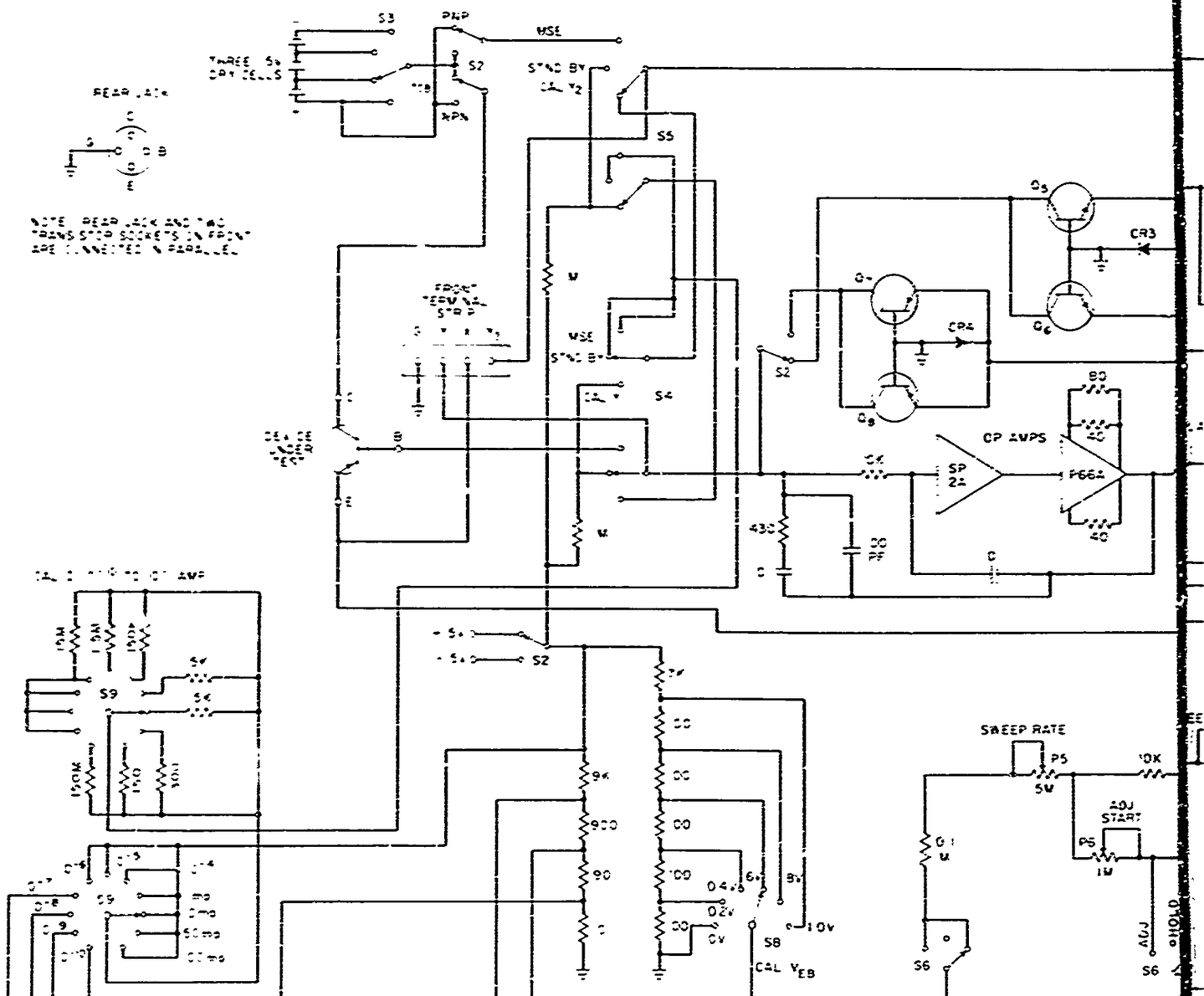
S6 is returned to the Adjust Start position, the curve is swept back towards the origin at a rate controlled by the value of P6. If switch S2 is thrown to the left, negative-going voltage ramp is generated which is suitable for testing NPN devices.

e. Slope Determination

If a low-frequency fixed incremental voltage, ΔV, is added to any given value of V_{EB}, it is possible, through the use of a phase-sensitive detector, to determine the slope of a log I_C or log I_B curve at that value of V_{EB}. As shown in Figure 59, a step-down transformer has been included to provide a low impedance source of low frequency alternating potential in series with the output of the V_{EB} generator.

f. Complete Circuit

The complete circuit of the curve tracer is shown in Figure 60. Included in the cabinet of the curve tracer are two the log conversion circuits shown in Figure 57, the V_{EB} sweep generator of Figure 59, regulated power supplies for the operational amplifiers and batteries to supply V_{CB} to the transistor under test. Built-in calibrating and switching facilities contribute to permanent accuracy and to the usefulness and convenience of the curve tracer.


S3
PNP
NSE
THREE 5½ DRY CELLS
S2
STAND BY
CAL V2
S5
M
FRONT PANEL STAB
NSE
STAND BY
S4
S2
Q5
CR3
Q4
CR4
Q6
Q8
80
40
OP AMPS
SP 22
P66A
40
430
50 PF
SWEEP RATE
P5
5W
10K
ADJ START
P5
1M
0.1 M
ADJ
HOLD
S6
S6
S9
150K
15K
5K
9K
900
90
0
100
0.4V
0.2V
0V
1.0V
S8
CAL VEB

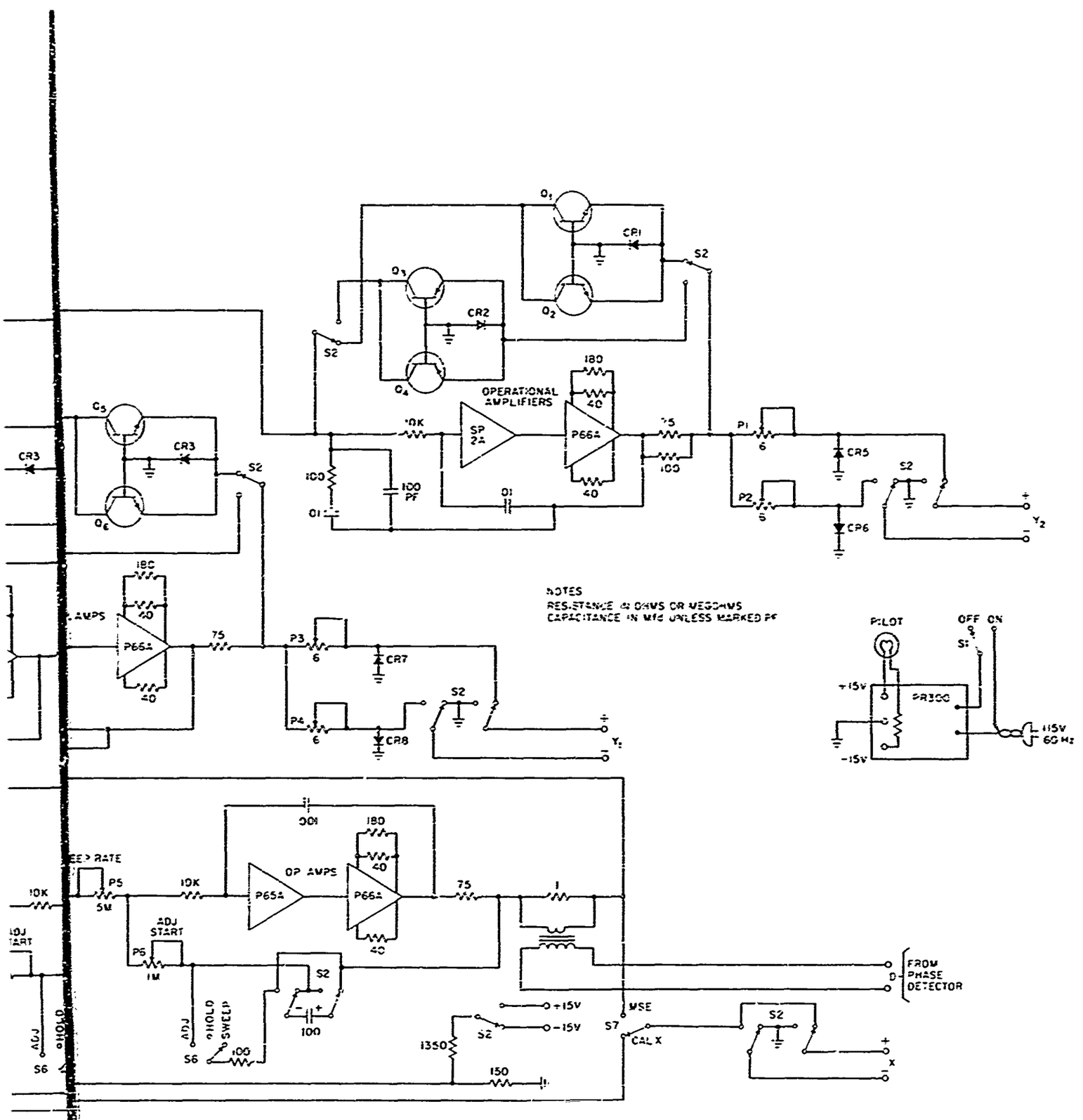


Figure 60. Complete Circuit, Logarithmic Curve Tracer

g. Calibration

Accurate voltages from 0 to 1.0 volt in 0.2-volt steps are available from switch S8 for calibrating the X or V_{EB} scale of the X-Y recorder.

Similarly, known currents are supplied from switch S9 through S4 or S5 to the log converters for calibrating either of the two Y scales of the recorder. These currents are provided in decade steps from 10^{-10} to 10^{-1} ampere.

h. External Connections

Transistors may be tested in either of the two front panel sockets or remotely through wires from a four pin Amphenol jack on the back of the set.

A terminal strip on the front panel provides direct access to the inputs of the Y1 and Y2 log converters and to the X channel output. Use of the terminal strip eliminates any small voltage drops which may occur across series switches at the higher currents. It is useful in plotting the V-I characteristics of high current diodes in the range from 10 to 100 ma.

13. CONCLUSIONS

The effects of electron and gamma irradiation on pair spectra in GaP have been studied. It was shown that damage defects are very effective in killing pair fluorescence. The experiments indicated that a non-radiative recombination mechanism was operative at deep damage defects but at the time the experiments were performed no satisfactory recombination mechanism had been formulated. This problem led to investigations of recombination at two other types of chemical defects. The first involved study of the kinetics of exciton recombination at isoelectronic traps such as nitrogen substituted for phosphorus in GaP. At these isoelectronic traps electrons and holes were found to annihilate each other with a high probability for photon production. From these studies it was possible to determine the nature of the traps (i.e., whether they are hole or electron traps) their oscillator strengths, capture cross sections, activation energies, and densities.

Studies of the kinetics of bound exciton recombination at neutral donors rather than isoelectronic centers revealed a more complex Auger-type mechanism. In this process, the exciton recombination energy is transferred to the weakly bound donor electron, which is ejected into the conduction band. This non-radiative process was found to predominate greatly over the radiative decay processes for these neutral centers. The discovery of this mechanism provided a possible explanation for the form of the killer action of deep damage center in GaP. Thus it is very probable that the hole, weakly bound at a Si-S pair, will transfer to a deep defect level with

acceptor characteristics. This neutral center can then act as a level for efficient non-radiative recombination. It appears that in the study of radiation-induced defects the possibility of this type of Auger recombination has seldom been considered. It is felt that the experimental techniques described will be valuable in further studies of recombination mechanisms at radiation defects. The recent discovery of fluorescence associated with recombination at the A center in silicon provides much impetus for continuing research in the present direction.

ESR studies of the SiG8 (E) center in silicon, heavily doped with phosphorus, show that the slow growth of the center with irradiation is probably, as previously assumed, caused by the gradual fall in the bulk Fermi level which is ultimately controlled by the SiG8 (E) center energy levels. Production rates of the SiG8 (E) centers are lower in LOPEX silicon than in comparable float-zone silicon. A previously unreported isotropic radiation damage center at g = 2.0045 is observed in most electron-irradiated, phosphorus-doped LOPEX crystals.

Following the summary of surface effects submitted in Scientific Report No. 1, an experimental program was initiated to investigate the effects of bias, temperature, and dose rate on the space charge accumulation process in oxidized silicon devices. It was found that the accumulation increases with bias of both polarities, decreases as the temperature increases, and is virtually independent of the dose rate. The results of these investigations indicates that the space charge accumulates when holes generated by the radiation are trapped while the corresponding electrons leave the oxide. An analysis based on this model shows how the space charge accumulates with dose and includes the effect of bias as a parameter. Comparison of the predicted dependence with that actually observed yields a value for the product of the mobility and lifetime of free electrons in the SiO_2 layer that is consistent with the assumption made in deriving the expressions for $\Delta V(t)$. The trapping model thus appears to be well established and well understood. However, the source of the hole traps has not yet been identified. If more is to be learned about these traps, new techniques such as thermoluminescence will be required. Once the origin of the hole traps is known, it may be possible to produce oxides, and hence devices, which are less sensitive to ionizing radiation.

Thermoluminescence has been applied to study the trapping process in oxidized silicon surfaces. Considerable differences were observed in the integrated luminescent intensity from three types of oxide-dry, wet, and deposited. Sodium was found to change the thermoluminescent characteristics significantly, but it does not seem to have been an initial contaminant in any of the oxides. Shallow hole traps apparently control the thermoluminescent behavior as they also do the behavior of irradiated MOS devices. Thermoluminescence studies in conjunction

with MOS capacitance studies could provide some real insight into both the radiation damage and reliability problems in oxidized silicon surfaces.

Co^{60}-gamma irradiations of Schottky Barrier diodes could not be used to study changes in interface state density because the radiation response of the devices was controlled by charge buildup at the SiO_2-Si-metal common interface. This charge buildup causes breakdown and/or tunneling to take place between the metal and semiconductor and can result in severe device degradation.

Transient recovery in silicon devices after fast neutron irradiation shows strong evidence of an electronic rather than an atomic process. This electronic process is attributed to the slow charge equilibration of large damage clusters (similar to dislocations). Such clusters would be minority carrier traps rather than recombination centers. Transient recovery of reverse bias junction capacitance after neutron irradiation has also been observed. An adequate understanding of the transient recovery phenomena will require a clearer experimental picture of the microscopic nature of neutron damage.

High-injection-level effects in neutron irradiated silicon power transistors and PNPN devices have been analyzed and compared with experimental results. PNPN devices have been shown to have power switching capability an order of magnitude greater than that of transistors exposed to the same neutron environment. Design criteria are given for radiation hardened PNPN and PIN devices. A PNPN device structure is proposed which should be capable of switching a 100 watt volt-ampere product at integrated fast neutron fluxes up to 10^{16} N cm^2.

A logarithmic curve tracer was designed and built which is capable of automatically recording both log I_c and log I_b against the emitter-base voltage of a transistor over more than eight decades of collector current (I_c) or base current (I_b). This curve tracer has been very useful in separating out the various recombination currents in irradiated devices.

14. CONTRIBUTORS

The authors of this report and contributors to the research reported herein, are as follows:

List of Authors

D. K. Wilson
J. P. Mitchell
J. D. Cuthbert
R. P. Blair

List of Contributors

D. K. Wilson	W. Jung
J. P. Mitchell	P. N. Tortoriello
J. D. Cuthbert	R. F. Mullholand
R. R. Blair	. A. Breitung
H. S. Lee	W. P. Knox

15. TALKS AND PUBLICATIONS

The talks and publications listed below resulted, wholly or partially, from work done under this control.

a. Talks

1. J. D. Cuthbert, "Effects of Radiation Damage on GaP Electroluminescence." Presented at A. P. S. Meeting, New York, June 1965. [Bull. Am. Phys. Soc., Vol. 10, No. 5 (1965) 595.]
2. J.D. Cuthbert and D.G. Thomas, "Fluorescent Decay Times of Excitons Bound to Isoelectronic Traps in GaP and ZnTe." To be presented at A.P.S. Meeting, New York, January 1967.
3. J.P. Mitchell, "Dose Rate Dependence of Surface Radiation Damage Effects in Planar Silicon Devices." Presented at NEREM 66 Conference, Boston, Massachusetts, 2, 3, 4 November 1966.
4. H.S. Lee and D K. Wilson, "Permanent Damage Radiation Effects in Narrow Base PNPN Devices." Presented at IEEE Conference on Nuclear and Space Radiation Effects, Stanford University, July 18-21, 1966.

b. Publications

1. D. F. Nelson, J. D. Cuthbert, P. J. Dean, and D. G. Thomas, "Auger Recombination of Excitons Bound to Neutral Donors in GaP and Si." Phys. Rev. Letters, Vol. 17 (1966) pp. 1262-1265.
2. J.D. Cuthbert and D.G. Thomas, "Fluorescent Decay Times of Excitons Bound to Isoelectronic Traps in GaP and ZnTe." Accepted for publication in Phys. Rev.
3. J.P. Mitchell and D.K. Wilson, "Surface Effects of Radiation on Semiconductor Devices," B.S.T.J., Vol. 46 (January 1967) 1-80.
4. D.K. Wilson and H.S. Lee, "Permanent Radiation Damage Effects in Narrow Base PNPN Devices." To be published.

16. REFERENCES

a. Section I

1. J. D. Cuthbert, "Studies of Recombination Fluorescence in Irradiated GaP and Unirradiated GaP, ZnTe, and Si," Scientific Report No. 2, Contract No. AF-19(628)4157, Bell Telephone Laboratories, Inc. (To be published)
2. D. G. Thomas, M. Gershenzon, and F. A. Trumbore, "Pair Spectra and 'Edge' Emission in Gallium Phosphide," Phys. Rev., Vol. 133, No. 1A (6 Jan 1964) A269.
3. O. S. Oen and D. K. Holmes, "Cross Sections for Atomic Displacements in Solids by Gamma Rays," J. Appl. Phys., Vol. 30, No. 8 (1959) 1289.
4. J. D. Cuthbert and D. G. Thomas (Accepted for publication in Phys. Rev.)
5. J. D. Cuthbert and D. G. Thomas, "Fluorescent Decay Times of Excitons Bound to Isoelectronic Traps in GaP and ZnTe," Bull Am. Phys. Soc., Vol. 12, No. 1 (January 1967) 101.
6. J. J. Hopfield, D. G. Thomas, and R. T. Lynch, "Isoelectronic Donors and Acceptors," Phys. Rev. Letters, Vol. 17, No. 6 (8 Aug 1966) 312.
7. J. D. Cuthbert and D. G. Thomas (to be published)
8. D. F. Nelson, J. D. Cuthbert, P. J. Dean, and D. G. Thomas, "Auger Recombination of Excitons Bound to Neutral Donors in Gallium Phosphide and Silicon," Phys. Rev. Letters, Vol. 17, No. 25 (19 Dec 1966) 1262.
9. P. J. Dean, J. D. Cuthbert, R. T. Lynch, and D. G. Thomas (To be published)
10. Yu. A. Kurskii, "The Possibility of Coherent Radiation from Localized Excitons in Silicon with Excitation Produced by Injection at a p-n Junction," Soviet Phys. Solid State, Vol. 7, No. 3 (Sept. 1965) 567.
11. G. D. Watkins, J. W. Corbett, and R. M. Walker, "Spin Resonance in Electron Irradiated Silicon," J. Appl. Phys., Vol. 30, No. 8 (1959) 1198.
12. G. D. Watkins and J. W. Corbett, "Defects in Irradiated Silicon: Electron Paramagnetic Resonance and Electron - Nuclear Double Resonance of the Si-E Center," Phys. Rev., Vol. 134, No. 5A (1 June 1964) A1359.
13. J. W. Corbett, "Electron Radiation Damage in Semiconductors and Metals," Solid State Physics, Suppl. 7, ed. F. Seitz and D. Turnbull (New York: Academic Press, 1966).
14. H. Saito and M. Hirata, "Nature of Radiation Defects in Silicon Single Crystals," Japan J. Appl. Phys., Vol. 2, No. 11 (Nov. 1963) 678.

15. E. Sonder and L. C. Templeton, "Gamma Irradiation of Silicon II. Levels in n-Type Float-Zone Material," J. Appl. Phys., Vol. 34, No. 11 (1963) 3295.

16. T. Tanaka and Y. Inuishi, "Hall Effect Measurement of Radiation Damage and Annealing in Si," J. Phys. Soc. Japan, Vol. 19, No. 2 (Feb. 1964) 167.

17. I. Chen and C. Kikuchi, "ESR of Neutron-Irradiated Silicon," Bull. Am. Phys. Soc., Vol. 10, No. 5 (1965) 582.

18. H. J. Stein, "Introduction Rate of Electrically Active Defects in n-Type Silicon by Nuclear Radiation," Sandia Report SC-R-65-938 (1965).

19. J. R. Carter, Jr., "Defect-Impurity Relationships in Electron-Damaged Silicon," IEEE Trans. on Nuclear Science, Vol. NS-13, No. 6 (December 1966) 24.

b. Section II

1. J. P. Mitchell and D. K. Wilson, "A Summary of Surface Effects of Radiation on Semiconductor Devices," Scientific Report No. 1, Contract No. AF-19(628)4157, Bell Telephone Laboratories, Inc. (1 Dec 1965).

2. J. P. Mitchell and D. K. Wilson, "Surface Effects of Radiation on Semiconductor Devices," B.S.T.J., Vol. 46 (January 1967) 1.

3. K. H. Zaininger, "Irradiation of MIS Capacitors with High Energy Electrons," IEEE Trans. on Nuclear Science, Vol. NS-13, No. 6 (December 1966) 237.

4. J. [illegible]olliday and E. J. Sternglass, "New Method for Range Measurements of [illegible]-Energy Electrons in Solids," J. Appl. Phys., Vol. 30, No. 9 (September 1959) 1428.

5. A. G. Stanley, "Comparison of MOS and Metal-Nitride-Semiconductor Insulated Gate Field-Effect Transistors Under Electron Radiation," IEEE Trans. on Nuclear Science, Vol. NS-13, No. 6 (December 1966) 248.

6. E. Kooi, "Influence of Heat Treatments and Ionizing Irradiations on the Charge Distribution and the Number of Surface States in the Si-SiO_2 System," IEEE Trans. on Electron Devices, Vol. ED-13, No. 12 (February 1966) 238.

7. A. S. Grove and E. H. Snow, "A Model for Radiation Damage in Metal-Oxide-Semiconductor Structures," Proc. IEEE, Vol. 54, No. 6 (June 1966) 894.

8. G. C. Messenger, E. J. Steele, and M. Neustadt, "Displacement Damage in MOS Transistors," IEEE Trans. on Nuclear Science, Vol. NS-12, No. 5 (October 1965) 78.

9. A. J. Speth and F. F. Fang, "Effects of Low-Energy Electron Irradiation on Si-Insulated Gate FETs," Appl. Phys. Letters, Vol. 7, No. 6 (15 September 1965) 145.

10. A. M. Goodman, "Photoemission of Holes from Silicon into Silicon Dioxide," Phys. Rev., Vol. 152, No. 2 (9 December 1966) 780.

11. F. J. Walter and R. R. Boshart, "The Rectification Process at Metal-Silicon Surface Barriers," IEEE Trans. on Nuclear Science, Vol. NS-13, No. 3 (June 1966) 189.

12. O. Leistiko, Jr. and A. S. Grove, "Breakdown Voltage of Planar Silicon Junctions," Solid State Electron, Vol. 9, No. 9 (September 1966) 847.

c. Section III

1. H. H. Sander and B. L. Gregory, "Transient Annealing in Semiconductor Devices Following Pulsed Neutron Irradiation," IEEE Trans. on Nuclear Science, Vol. NS-13, No. 6 (December 1966) 53.

2. D. Binder and D. T. Butcher, "Rapid Annealing of Silicon Transistors," Hughes Aircraft Co. Tech. Rpt. No. AFWL-TR-66-145 (February 1967).

3. J. Nihoul and L. Stals, "On the Reaction Order of Bimolecular Recovery Processes in Solids," Phys. Stat. Sol., Vol. 17, No. 1 (1966) 295.

4. J. R. Parsons, R. W. Bailuffi, and J. S. Koehler, "Direct Observation of Neutron Damage in Germanium," Appl. Phys. Letters, Vol. 1, No. 3 (1 November 1962) 57.

5. M. Bertolotti, T. Papa, D. Sette, and G. Vitale, "Electron Microscope Observation of High-Energy-Neutron-Irradiated Germanium," J. Appl. Phys., Vol. 36, No. 11 (November 1965) 3506.

6. P. L. F. Hemment and E. M. Gunnersen, "Transmission Electron Microscopy of Fast-Neutron-Irradiated Silicon," J. Appl. Phys., Vol. 37, No. 7 (June 1966) 2912.

7. Wun Jung and G. S. Newell, "Spin-1 Centers in Neutron-Irradiated Silicon," Phys. Rev., Vol. 132, No. 2 (15 October 1963) 648.

8. J. W. Corbett, "Electron Radiation Damage in Semiconductors and Metals," Solid State Physics, ed. F. Seitz and D. Turnbull, Vol. 18, Suppl. 7 (New York: Academic Press, 1966).

9. R. E. Whan, "Oxygen-Defect Complexes in Neutron-Irradiated Silicon," J. Appl. Phys., Vol. 37, No. 9 (August 1966) 3378.

10. O. L. Curtis, Jr., "Effects of Oxygen and Dopant on Lifetime in Neutron-Irradiated Silicon," IEEE Trans. on Nuclear Science, Vol. NS-13, No. 6 (December 1966) 33.

11. H. J. Stein, "Introduction Rate of Electrically Active Defects in n-type Silicon by Nuclear Radiation," Sandia Report SC-R-65-938 (1965).

12. H. J. Stein, "Comparison of Neutron and Gamma-Ray Damage in n-type Silicon," J. Appl. Phys., Vol. 37, No. 9 (August 1966) 3382.

13. B. R. Gossick, "Disordered Regions in Semiconductors Bombarded by Fast Neutrons," J. Appl. Phys., Vol. 30, No. 8 (August 1959) 1214.

14. V. Heine, "Dangling Bonds and Dislocations in Semiconductors," Phys. Rev., Vol. 146, No. 2 (10 June 1966) 568.

15. G. D. Watkins and J. W. Corbett, "Defects in Irradiated Silicon: Electron Paramagnetic Resonance of the Divacancy," Phys. Rev., Vol. 138, No. 2A (19 April 1965) A543.

16. Y. Inuishi and K. Matsuura, "Radiation Damage and Annealing of Carrier Life-Time in Silicon," J. Phys. Soc. Japan, Vol. 13, Suppl. III (March 1963) 240.

17. J. J. Wysocki, "Thermally Stimulated Conductivity of Neutron-Irradiated Silicon," J. Appl. Phys., Vol. 36, No. 9 (September 1965) 2968.

18. K. D. Glinchuk, N. M. Litovchenko, and V. A. Nevikova, "Trapping of Carriers in Plastically Deformed Silicon," Soviet Phys. Solid State, Vol. 8, No. 3 (September 1966) 777.

19. C. T. Kirk, Jr., "A Theory of Transistor Cutoff Frequency (f_T) Falloff at High Current Densities," IRE Trans. on Electron Devices, Vol. ED-9, No. 2 (March 1962) 164.

APPENDIX A

AUTOMATIC C-V_G CURVE PLOTTER

A convenient method of monitoring the positive charge accumulation in the SiO_2 layer of a MOS-FET is the use of the capacitance-voltage characteristics of the gate-to-substrate capacitance of the device. The tedious and time-consuming chore of point-by-point plotting of the numerous C-V curves required dictated the need for a much improved technique - preferably automatic. A search of commercially available equipment led to the acquisition of a Model 71A Boonton L-C meter. The L-C meter provides a DC analog output of the capacitance as read on the meter which is capable of driving directly any high input impedance recorder and has provision for supplying bias to the "test" capacitor from an external source. All that was required then was an X-Y recorder, which was readily available, and a suitable bias supply capable of sweeping over the desired voltage ranges at a sufficiently slow rate to allow equilibrium of charge in the "fast" interface states.

Since a sweep generator meeting the specific requirements of this study was not available, one had to be constructed. A schematic of the circuit designed for this purpose is shown in Figure A-1. The power supply is capable of providing independently +100 or -100 volts maximum in 10-volt steps, with one additional 5-volt range at the low end. The heart of the sweep generator is a 10-turn, motor-driven Helipot, with a special center-tap, that sweeps the entire voltage span from plus volts through zero to minus volts or vice versa in one minute. The Helipot is easily changed should longer or shorter sweep intervals be desired. Since the speed of the motor-driven pot is voltage sensitive, a well regulated supply was incorporated to assure a constant turning speed. The overall measuring technique is shown in block diagram in Figure A-2.

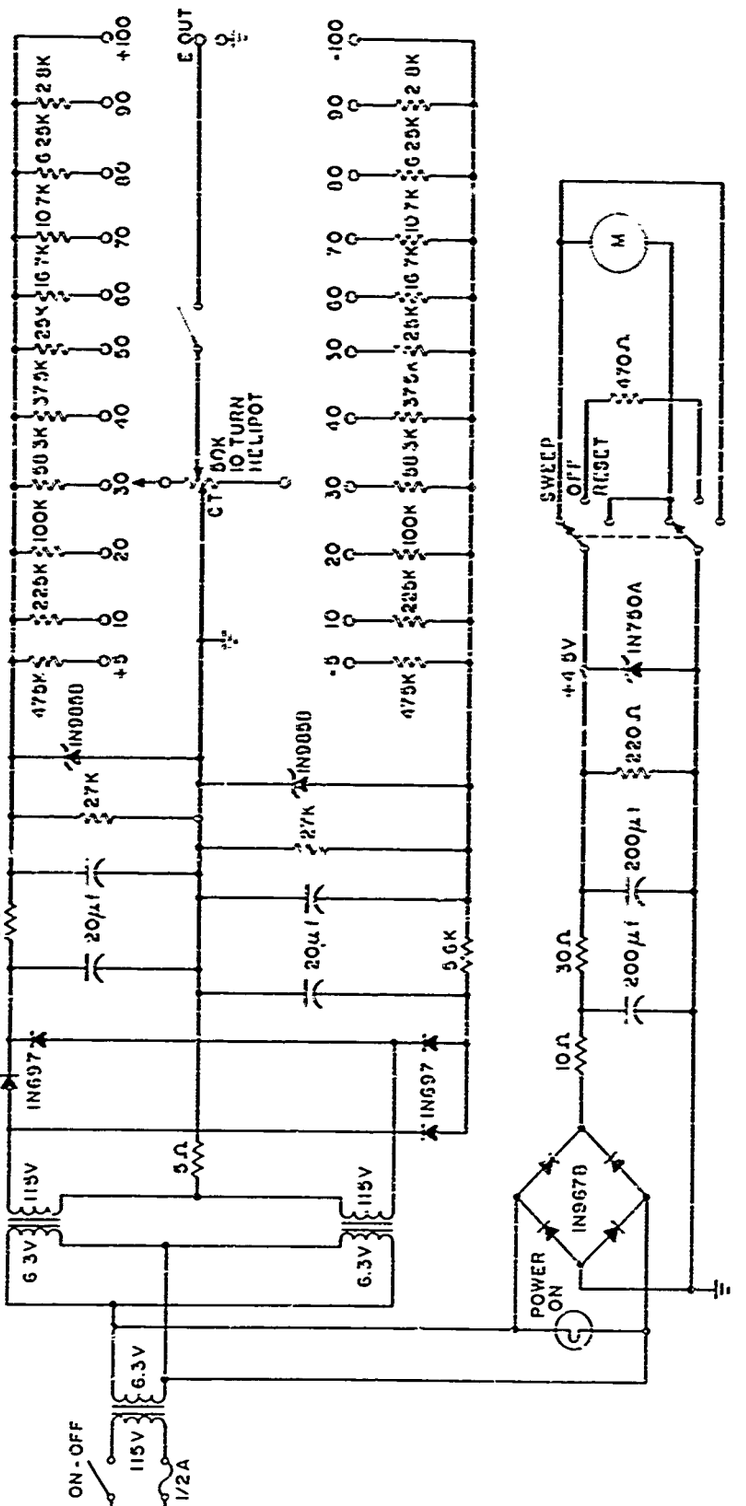


Figure A-1. Schematic of Sweep Generator Circuit

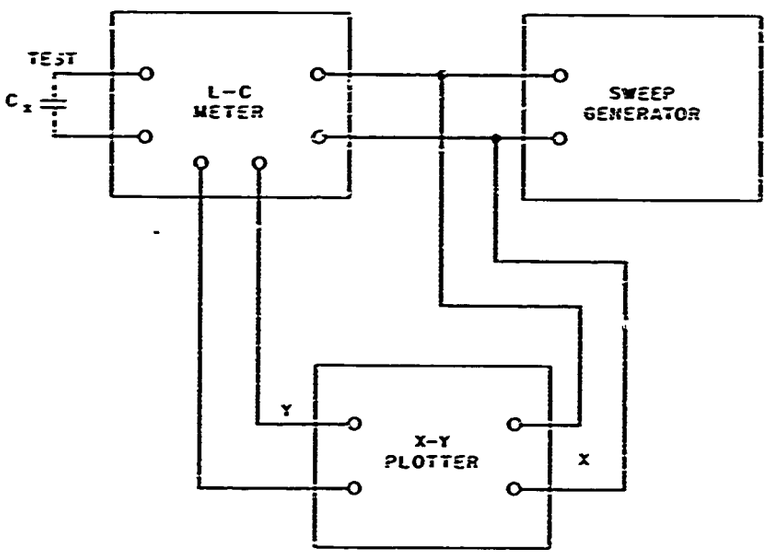


Figure A-2. Block Diagram of Automatic Plotting System

APPENDIX B

INITIAL EQUILIBRIUM ELECTRON DENSITY

Consider the situation shown in Figure 19. The electric field in the oxide, E, is uniform throughout the oxide and of magnitude V_G/x_o. It is assumed that electrons are generated at a rate of g/cm^3-s. The electrons have a lifetime τ_e and a mobility μ. At any point x in the oxide the electron density, n, is given by the continuity equation

$$\frac{\partial n}{\partial t} = \left(g - \frac{n}{\tau_e}\right) + \mu \frac{\partial \Xi n}{\partial x} + D \frac{\partial^2 n}{\partial x^2} \tag{1}$$

where D is the diffusion constant for the electrons.

It will be assumed that diffusion is negligible compared to drift, i.e., $D \sim 0$, and that no space charge accumulates in the time required for n to reach equilibrium, $(\partial n / \partial t) = 0$. Equation 1 becomes

$$\mu E \frac{dn}{dx} - \frac{n}{\tau_e} = -g, \quad E = -\frac{V_G}{X_0} \tag{2}$$

Since the Si cannot supply electrons to the oxide, n(o) = 0. Using this boundary condition, the solution to equation 2 is easily found to be

$$n(x) = g\tau_e\left[1 - \exp\left(-a\frac{x}{x_o}\right)\right] \tag{3}$$

where $a = \dfrac{x_o^2}{\mu \tau_e V_G}$

APPENDIX C

ELECTRIC FIELD IN THE OXIDE AT $X = X_O$ AT ANY TIME

Figure 22 illustrates the situation in the MOS structure at any time t. The purpose of this Appendix is to derive an expression for the field E in the oxide at $x = x_o$ in terms of the space charge $Q_{ss}(t)$. E is uniform in the region $d \leq x \leq x_o$ since there is no space charge in the region. In the region $o \leq x \leq d$ Poisson's equation yields

$$\frac{d^2V}{dx^2} = - \frac{qP_T(t)}{\epsilon \epsilon_o} \tag{1}$$

where V is the potential at any point x, ϵ the dielectric constant of SiO_2, and ϵ_o the permittivity of free space. Integration of equation (1) gives

$$\frac{dV}{dx} = - \frac{qP_T(t)}{\epsilon \epsilon_o} x + A$$

where $A = \left.\frac{dV}{dx}\right|_{x = o} = - E(o)$

With the aid of Gauss' law and the relation $Q_{Si} = - \left[Q_{ss}(t) + Q_m(t)\right]$ it is easily shown that

$$E(o) = - \frac{Q_{ss}(t) + Q_m(t)}{\epsilon \epsilon_o} \tag{2}$$

Now

$$Q_m(t) = C_{ox} V_G - Q_{ss}(t) \frac{d}{2x_o} \tag{3}$$

and

$$Q_{ss}(t) = qP_T(t)d \tag{4}$$

$$C_{ox} = \frac{\epsilon \epsilon_o}{x_o} \tag{5}$$

Combining equations (2), (3), (4), (5) gives

$$A = \frac{qP_T(t)d}{\epsilon \epsilon_o} \left(1 - \frac{d}{2x_o}\right) + \frac{V_G}{x_o} \tag{6}$$

Hence

$$E(x_o) = E(d) = - \left.\frac{dV}{dx}\right|_{x = d}$$

$$= \frac{qP_T(t)d}{\epsilon\epsilon_o} - A \qquad (7)$$

$$E(x_o) = - \frac{V_G}{x_o} + \frac{qP_T(t)d^2}{2\epsilon\epsilon_o x_o}$$

ATTACHMENT I

METAL-OXIDE-SEMICONDUCTOR (MOS) CAPACITORS

1. INTRODUCTION

Semiconductor devices employing SiO_2, such as passivated Si planar transistors, monolithic integrated circuits, or metal-oxide-semiconductor (MOS) field-effect transistors (FET) are known to suffer degradation when operated at elevated temperatures or exposed to radiation. The SiO_2 and the SiO_2-Si interface regions of these devices apparently hold the key to the degradation process. By some means, not at present understood, positive charge accumulates in the oxide at the SiO_2-Si interface and causes degradation either through the formation of channels on the device surface or by increasing the surface carrier generation-recombination.

The MOS capacitor is being used increasingly in the study of SiO_2 and the SiO_2-Si interface.[1-5] These capacitors are simple to construct and, through relatively straightforward experiments, can yield useful information about charge processes in the oxide. The MOS capacitor is, therefore, likely to be important in the study of radiation effects on the surfaces of semiconductor devices. This attachment is intended to describe the construction, theory, and uses of these capacitors.

2. THE MOS STRUCTURE

A MOS capacitor is usually formed by thermally growing a few thousand angstroms of oxide on a p- or n-type Si wafer a few mils thick, and then depositing a metal electrode, the gate, on the SiO_2 surface. The oxide is removed from the bottom surface of the wafer so that an ohmic contact can be made to the Si. Electrical contact to the gate is often made by means of a pressure contact. The resulting structure is a parallel plate capacitor with the gate and surface region of the silicon acting as the plates and the oxide as the dielectric between them.

3. MOS CAPACITANCE AS A FUNCTION OF APPLIED VOLTAGE

a. Qualitative Discussion

Since a semiconductor forms one plate of a MOS capacitor, the capacitance will be dependent on the applied voltage. The gross features of this voltage dependence can be predicted from a qualitative argument. The surface layer of the Si will be accumulated, depleted, or inverted depending on the gate potential. The charges stored in the capacitor for these three conditions are illustrated in Figure I-1 for p-type silicon in the absence of trapped charge in the SiO_2. Also shown in Figure I-1 are the corresponding band structures at the Si surface.

For $V_G < 0$, (a) in Figure I-1, a charge, Q_S, comprised of mobile holes equal in magnitude to the negative charge on the gate, Q_G, collects in the Si surface layer. For this condition, the MOS capacitance plates are essentially separated by the oxide. The capacity (per unit area) is, therefore, that due to the oxide,

$$C_{ox} = \frac{K_s \epsilon_o}{X_o}$$

where K_s and X_o are the dielectric constant and thickness of the oxide respectively and ϵ_o is the permittivity of free space.

For $V_G > 0$, (b) in Figure I-1, the hole concentration in the Si surface becomes depleted and the charge density in the surface layer, Q_S, comes from exposed, fixed acceptor impurities (i.e., $Q_S = -qN_A$, where N_A is the density of acceptors). The depleted surface of the Si acts as a dielectric and hence the separation of the capacitor plates is increased causing a decrease in the capacity. When $V_G >> 0$, (c) in Figure I-1, electrons are attracted into the Si surface layer and the surface becomes inverted. Q_S is now comprised of a negative charge of mobile electrons, Q_n, and charged acceptor atoms, $Q_S = Q_n - qN_A$. It will be shown later that as Q_S increases, $Q_n - Q_S$, i.e., the negative charge in the Si surface is made up almost entirely of mobile minority carriers. Thus, as the surface becomes more inverted, the depletion region stops its advance into the Si and sufficient minority carriers collect at the Si surface so that the capacity again becomes essentially that due to the oxide, C_{ox}. The expected variation of the MOS capacity, C, normalized to C_{ox} is shown in the low-frequency curve of Figure I-2. For a MOS capacitor with n-type Si the capacity minimum will, of course, occur at negative values of gate voltage. Trapped charge present in the oxide acts as a bias on the capacitor and causes the curve to shift parallel to the abscissa, in the positive direction for negative charge and in the negative direction for positive charge.

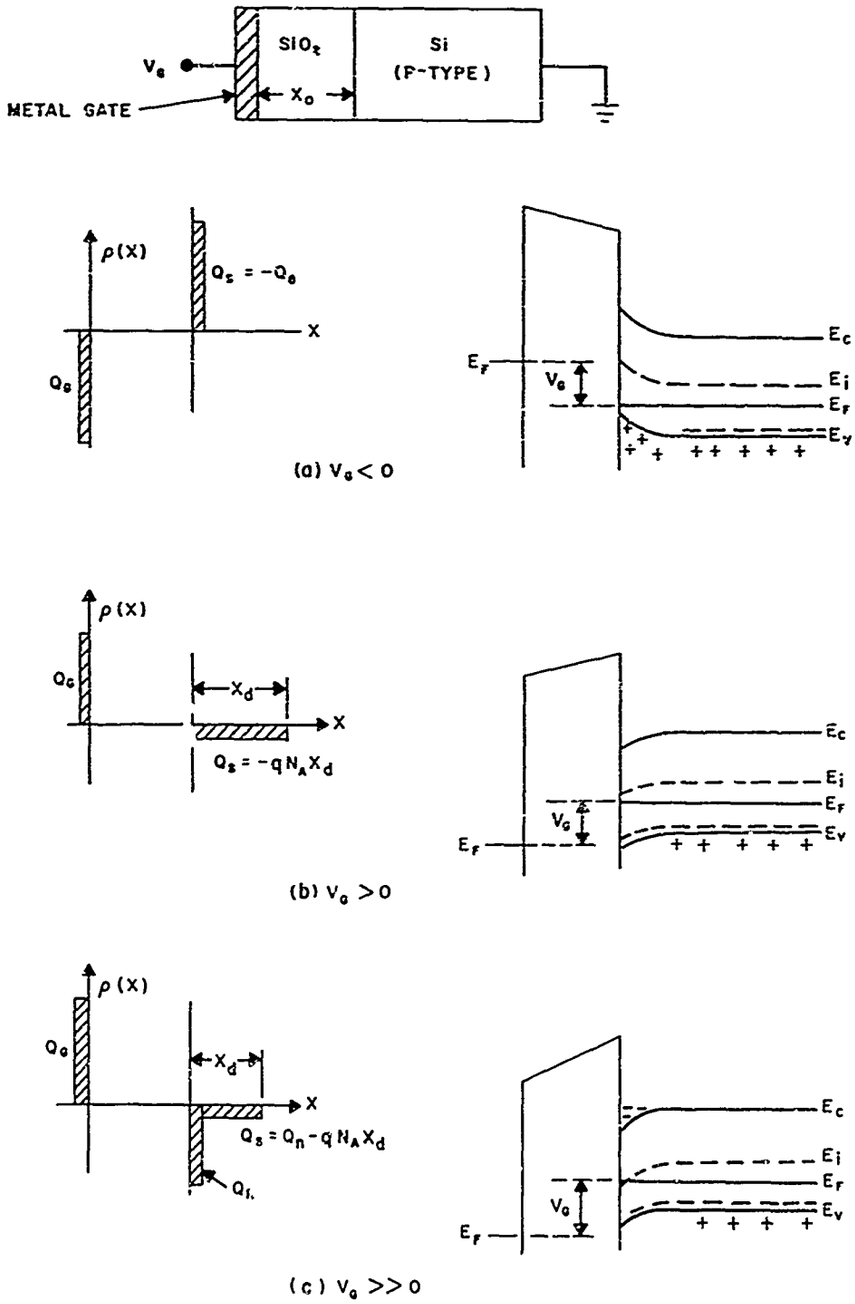


Figure I-1. Charge Distribution and Band Structure at the Si Surface for a P-Type Si MOS Capacitor

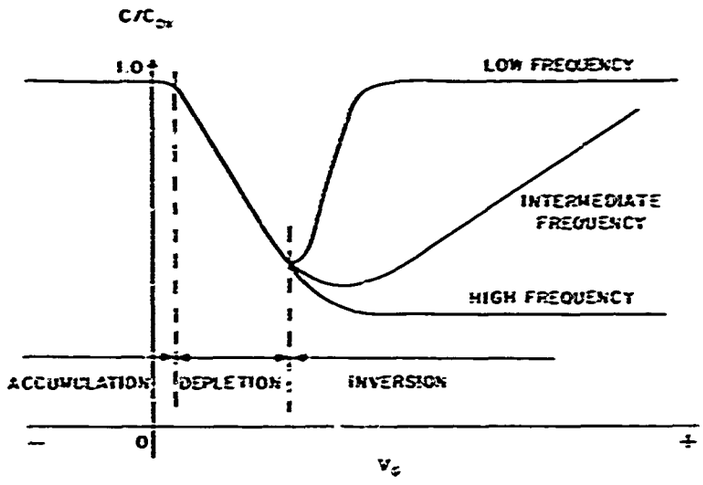


Figure I-2. Variation of C/C_{ox} with V_G for a P-Type Si MOS Capacitor at Various Measuring Frequencies

b. Frequency Dependence

The procedure for measuring MOS capacitance involves applying a known d-c bias to the capacitor and then measuring the capacitance with a small, superimposed a-c signal. In the discussion above it was tacitly assumed that the mobile carriers in the semiconductor surface were in equilibrium at all times, i.e., that variations of the potential difference across the capacitor were slow enough for the carrier concentration to follow. Majority carriers can remain in equilibrium with a-c signals for frequencies up to $\sim 10^9$ Hz. Hence, for the Si surface in the accumulation or depletion condition, the surface is controlled by majority carriers and the capacity is virtually independent of frequency.

For the inversion condition, however, minority carriers control the surface region. The time constants involved in the recombination and generation of minority carriers are quite long and hence the ability of the minority carrier density in the inversion layer to follow the a-c signal will depend on the signal frequency. At low frequencies ($\lesssim 10^2$ Hz) the minority carrier density can follow the a-c signal, and the capacitance variation with applied voltage will follow the low-frequency curve in Figure I-2. At high frequencies ($\gtrsim 10^4$ Hz), on the other hand, the surface density of minority carriers cannot follow the rapid a-c voltage variations and the measured capacity will be due only to majority carriers in the depletion region.

The (differential) capacity which is proportional to the change of the amount of charge in the Si, i.e., $C = dQ_s/dV$, will, therefore, remain at a lower value at high frequencies because dQ_s will be due only to the majority carrier concentration changes. The variation of C in this case is illustrated by the high-frequency curve in Figure I-2. The capacity variation at an intermediate frequency ($\sim 10^3$ Hz) is also shown in Figure I-2.

c. Quantitative Analysis

A quantitative analysis of MOS capacitance variation with applied voltage has been given by several authors.[1-4,6] For the study of radiation effects, however, a simple version of the theory which gives the more gross features of capacitance variation will be sufficient. The analysis given below follows that given by Grove et al.[3,4]

Consider a p-type MOS capacitor, such as that illustrated in Figure I-1, under the following assumptions.

1. No surface states are present in the oxide, i.e., no charge is trapped in the oxide.
2. There is no work function difference between the metal gate and the silicon.
3. The d-c bias changes slowly enough to permit carrier equilibrium at all times.
4. The doping level in the Si surface layer is the same as that in the bulk Si.

For all Si surface conditions the total charge in the silicon, Q_s is given by

$$Q_s = \int_0^\infty \rho(x)\, dx$$

$$\rho(x) = q\left(p - n + N_D - N_A\right)$$

where

p = the hole concentration
n = the electron concentration
N_D = the density of donors
N_A = the density of acceptors
q = the magnitude of the electronic charge

The potential in the semiconductor, Φ, is given by Poisson's equation

$$\nabla^2\Phi = \frac{-\rho(x)}{K_S\epsilon_o}$$

where K_S is the dielectric constant of Si. Now

$$p = n_i \exp\left(u_F - u\right)$$
$$n = n_i \exp\left(u - u_F\right)$$

where

n_i = the intrinsic concentration of carriers
u = the potential in the semiconductor in units of kT (i.e., $u = q\Phi/kT$)
u_F = the Fermi energy level

The quantity $N_A - N_D$ can be found from the values of p and n in the bulk where $u = 0$ by definition.

$$N_A - N_D = [n - p]_{u=0}$$
$$= 2n_i \sinh u_F$$

Hence

$$\rho = -2n_i q\left[\sinh\left(u - u_F\right) + \sinh u_F\right]$$

Substituting the expression for ρ into Poisson's equation yields, after integrating once and using Gauss' Law,

$$\frac{du}{dx} = \pm\frac{\sqrt{2}}{L_D}\left[\cosh\left(u - u_F\right) - \cosh u_F + u \sinh u_F\right] \qquad (1)$$

and

$$Q_S = -2\frac{u_S}{|u_S|} L_D q n_i \sqrt{2}\left[\cosh\left(u_S - u_F\right) - \cosh u_F + u_S \sinh u_F\right]^{1/2} \qquad (2)$$

where

$$L_D^2 = \frac{kT}{q}\frac{K_S\epsilon_o}{2n_i q}$$

and $u_S = q\sigma_S/kT$ is the potential at the Si surface.

For the case of inversion

$$Q_n = -q\int_0^{x_i} n(x)dx = -q\int_{u_S}^{u_F}\frac{n(u)du}{(du/dx)}$$

where x_i is the depth of the surface inversion layer. Using equation (1), Q_n becomes

$$Q_n = \frac{-u_S}{\sqrt{2}\,|u_S|}\, qn_i L_D \int_{u_F}^{u_S} \frac{\exp\left(u - u_F\right) du}{\left[\cosh\left(u - u_F\right) - \cosh u_F - u \sinh u_F\right]^{1/2}} \qquad (3)$$

From equations (2) and (3), Q_S and Q_n can be calculated for a given temperature, applied voltage (u_S), and doping level (u_F). Figure I-3 shows $|Q_n/q|$ as a function of $|Q_S/q|$ at 300°K for several doping levels.[4] From these curves it is apparent that as the charge in the Si, Q_S, increases, the charge, Q_n, contained in the inversion layer also increases, and $Q_n - Q_S$ for large values of Q_S. Thus, at high inversion levels the depletion layer stops increasing in width and the negative charge on the semiconductor plate is comprised of minority carriers (electrons) at the SiO_2-Si interface. At low frequencies the capacity is thus C_{ox} for the inversion condition.

An effective depletion width, X_d, can be defined by the equation

$$Q_S = Q_n + q\left(N_D - N_A\right) X_d \qquad (4)$$

Figure I-4 shows the variation of X_d with $|Q_S/q|$ at 300°K.[4] As expected, X_d becomes almost constant at high $|Q_S/q|$ values.

d. The Effect of Surface States and Work Function Difference

In any real MOS structure there will be a work function difference, Φ_{MS}, between the metal gate and the Si as well as some charge trapped in the oxide, Q_{ss}. The band structure at the surface of a p-type semiconductor with no applied voltage but with the effects of the work function difference and trapped charge included is shown in Figure I-5a. The corresponding charge distribution is shown in Figure I-5b.

From Figure I-5a it is apparent that

$$\varphi_M + V_{oo} = \chi - \varphi_{so} + \frac{E_g}{2} + \varphi_F \qquad (5)$$

where

Φ_M is the work function of the gate metal
χ is the electron affinity of the Si
E_g is the energy gap of Si
V_{oo} and Φ_{so} are the potential drop across the oxide and the Si surface potential respectively at zero applied gate voltage

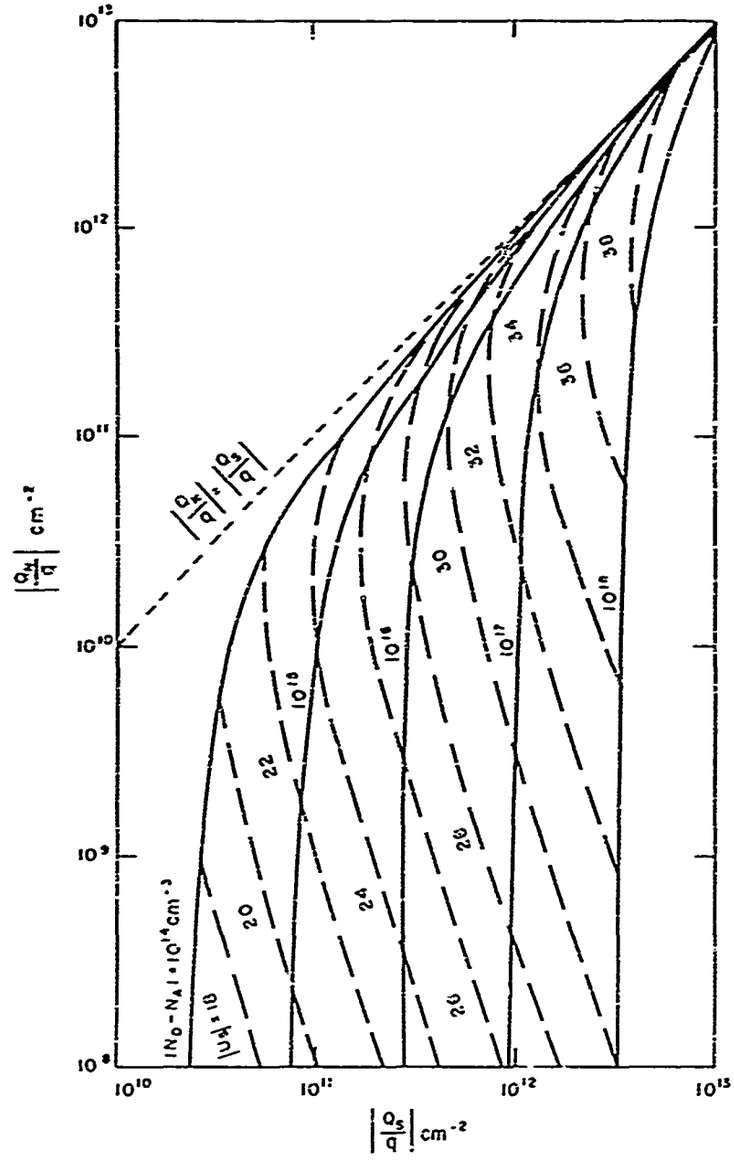


Figure I-3. $|Q_S/q|$ versus $|Q_n/q|$ for Various Si Doping Levels at 300°K

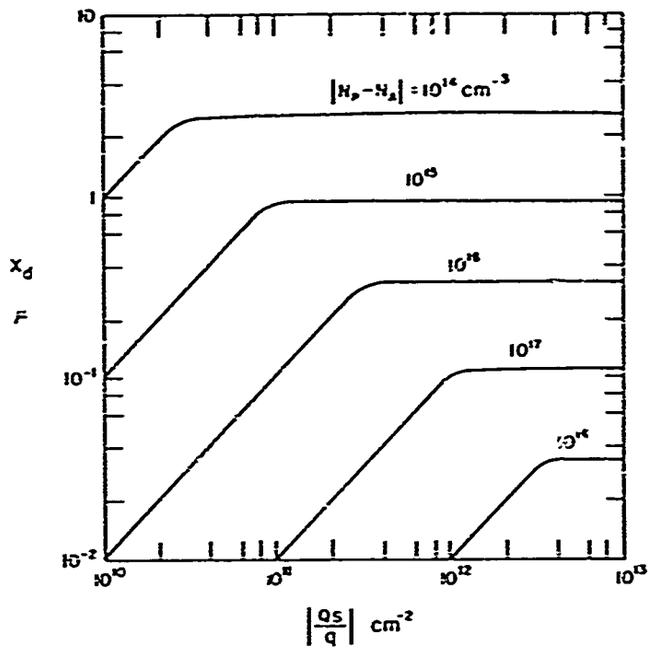


Figure I-4. Variation of X_d with $|Q_S/q|$ for Various Si Doping Levels at 300°K

By definition

$$\varphi_{Ms} = \varphi_M - \left(\chi + \frac{E_g}{2} + \varphi_F\right)$$

or using equation (5)

$$\varphi_{Ms} = -\left(\varphi_{so} + V_{oo}\right) \tag{6}$$

In the following treatment, no distinction will be made among the oxide surface states arising from various causes and all charge in these states will be treated as if it were located at the SiO_2-Si interface. Also, it will be assumed that the charge in these surface states, Q_{ss}, is infinitely slow, i.e., Q_{ss} is independent of the surface potential.

If a potential, V_G, is applied to the gate the Si surface band structure and MOS charge distribution will be altered to those shown in Figure I-6. For this situation

$$V_G = \left(V_o - V_{oo}\right) + \left(\varphi_s - \varphi_{so}\right)$$

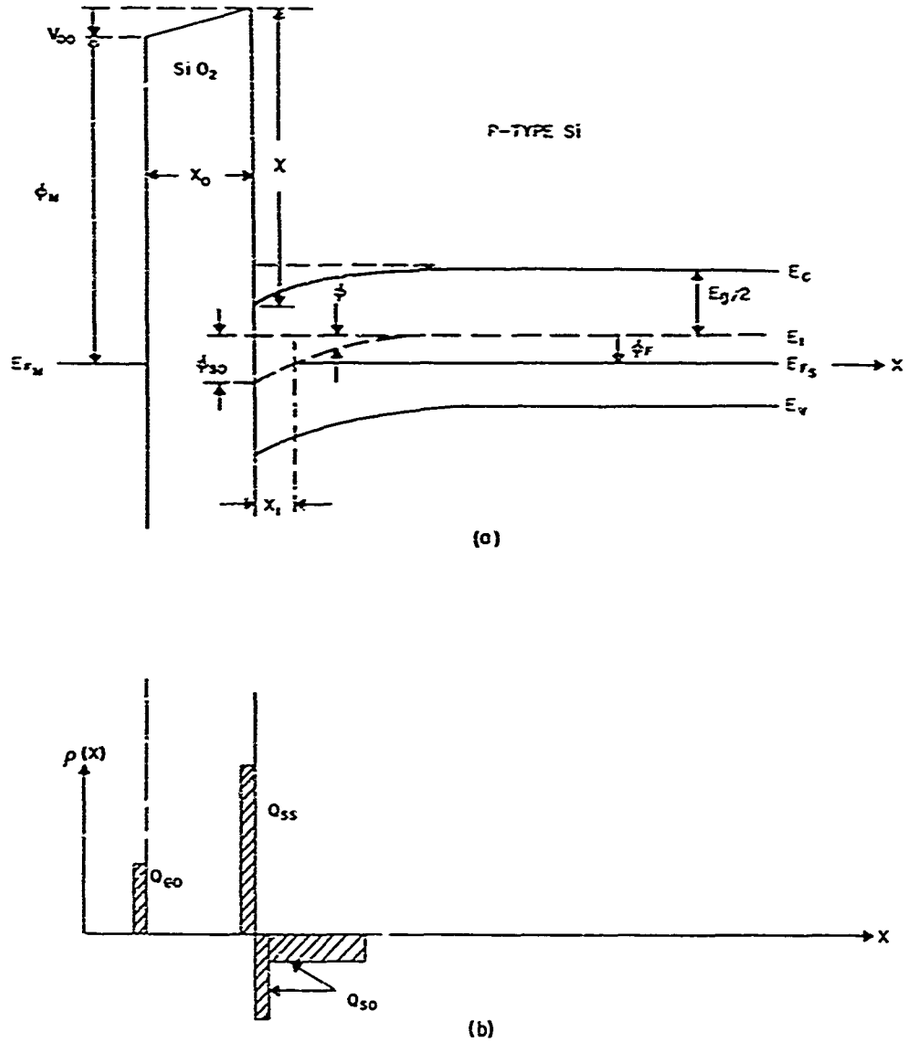


Figure I-5. Effect of a Work Function Difference and Trapped Charge on the Si Surface Band Structure for $V_G = 0$, and Corresponding Charge Distribution

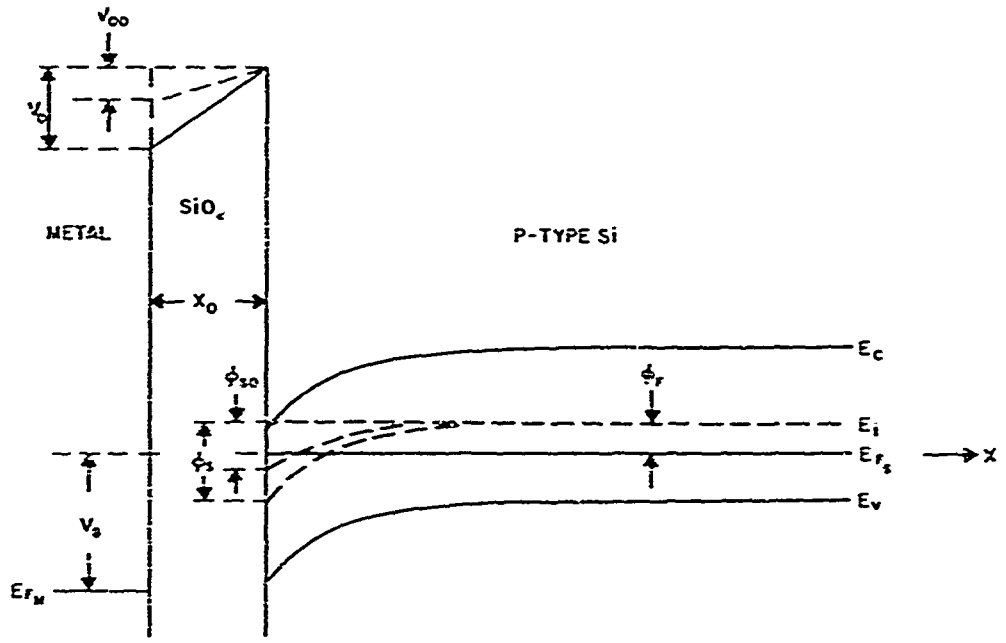


Figure I-6. Effect of a Work Function Difference and Trapped Charge on the Si Surface Band Structure for $V_G > 0$, and Corresponding Charge Distribution

or, using equation (6)

$$V_G = V_o + \varphi_s + \varphi_{Ms} \tag{7}$$

By charge neutrality requirements

$$Q_G + Q_{ss} + Q_s = 0 \tag{8}$$

Also $Q_G = V_o C_{ox}$ and hence combining equations (7) and (8)

$$V_G - \varphi_{Ms} + \frac{Q_{ss}}{C_{ox}} = \varphi_s - \frac{Q_s}{C_{ox}} \tag{9}$$

Now the total MOS capacitance, C, is defined by the relation

$$C = \frac{dQ_G}{dV_G}$$

But from equation (7)

$$dV_G = dV_o + d\Phi_s$$

Hence

$$C = \left(\frac{dV_o}{dQ_G} + \frac{d\varphi_s}{dQ_G}\right)^{-1}$$

$$C = \left[\frac{1}{C_{ox}} + \frac{1}{(C_s + C_{ss})}\right]^{-1} \tag{10}$$

where $C_s \equiv -dQ_s/d\Phi_s$ and $C_{ss} \equiv -dQ_{ss}/d\varphi_s$, i.e., a capacitance is associated with the charge in the Si, Q_s, and the charge in the oxide, Q_{ss}.

According to equation (10) the MOS structure can be pictured as a combination of three capacitors as shown in Figure I-7a. However, according to the assumption of infinitely slow states in the oxide, $dQ_{ss} = 0$ and hence $C_{ss} = 0$. The MOS equivalent circuit reduces to a series combination of C_{ox} and C_s as shown in Figure I-7b. There will, of course, be resistances associated with the bulk of the material in the MOS structure and leakage through the SiO_2 layer.

e. Low Frequency Capacity

At low frequencies both majority and minority carrier densities in the Si surface can follow the a-c signal and hence for all three surface conditions

$$C_s = -\frac{dQ_s}{d\varphi_s} = \frac{q}{kT}\frac{dQ_s}{du_s}$$

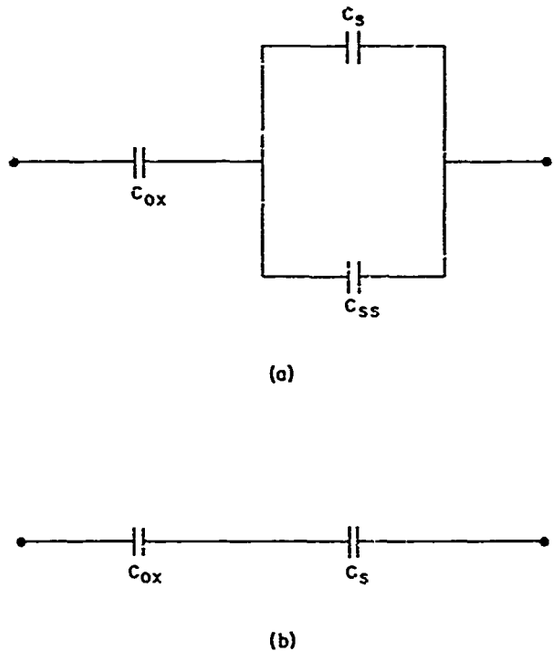


Figure I-7. MOS Equivalent Circuit (Resistances Omitted)

Substituting for Q_S from equation (2)

$$C_S = K_S \epsilon_0 q \frac{\left(p_S - n_S + N_D - N_A\right)}{Q_S} \tag{11}$$

where p_S and n_S are the hole and electron concentration respectively at the Si surface.

It is now possible to calculate C as a function of voltage. The procedure is straightforward although somewhat lengthy. The quantities C_{ox}, T, $N_A - N_D$, and φ_{Ms} are all determined in advance for a given MOS capacitor. The first step is to select an arbitrary value of φ_S (i.e., u_S) and calculate the corresponding value of Q_S from equation (2). Once Q_S is known $V_o - \Phi_{Ms} + Q_{ss}/C_o$ can be found from equation (9). Next, C_S is found from equation (11) and this value of C_S together with the known value of C_{ox} gives C through equation (10) ($C_{ss} = 0$). The calculation is repeated for other values of φ_S until a curve of C versus $V_o - \varphi_{Ms} + Q_{ss}/C_{ox}$ can be plotted.

f. High Frequency Capacity

The MOS capacity will be the same at high and low frequencies except when the Si surface becomes inverted. For inversion, Q_n is constant since only majority carriers can follow the a-c signal. Now $C_s = -(dQ_s/d\varphi_s)$ and using equation (4) results in

$$C_s = q(N_A - N_D)\frac{dX_d}{d\varphi_s}$$

Grove et al.[3] have shown that

$$\frac{dX_d}{d\varphi_s} = \frac{1}{q(N_A - N_D)}\frac{K_s\epsilon_o}{X_d}$$

and hence

$$C_s = \frac{K_s\epsilon_o}{X_d} \tag{12}$$

The procedure for calculating C as a function of $V_o - \varphi_{Ms} + Q_{ss}/C_{ox}$ is similar to that used in the low frequency case. From equation (3), Q_s is obtained for a particular value of φ_s. From Figure I-4 the corresponding value of X_d is determined and used to calculate C_s from equation (12). Now C_s and C_{ox} may be used to give C from equation (10). $V_o - \varphi_{Ms} + Q_{ss}/C_{ox}$ is determined from equation (9) and there is now one point on the graph. The calculations are repeated until sufficient points have been obtained.

4. DETERMINATION OF THE DENSITY OF TRAPPED CHARGE, Q_{ss}

In practice C is measured at some fixed frequency as a function of d-c bias, V_G. The calculation outlined above, however, gives C as a function of $V_G - \varphi_{Ms} + Q_{ss}/C_{ox}$. If the experimental and calculated curves are parallel then the difference in abscissa between them is due to the term $-\varphi_{Ms} + Q_{ss}/C_{ox}$. Figure I-8 shows the calculated and experimental curves for a p-type capacitor as reported by Grove et al.[4] If φ_{Ms} and C_{ox} are known Q_{ss} can be determined. Grove estimates φ_{Ms} to be -0.7V for p-type Si ($N_A \sim 10^{16}$ cm^{-3}). Q_{ss}/q is then found to be $\sim 2 \times 10^{11}$ cm^{-2}. Values of Q_{ss}/q reported by other authors vary but are generally in the range 10^{11} to 10^{12} cm^{-2}. The charge trapped in the oxide is found to be positive regardless of the conductivity type of the Si.

When the Si surface is inverted the MOS capacitance is strongly dependent upon the doping level of Si. Figure I-9 shows the variation C/C_{ox} with gate voltage

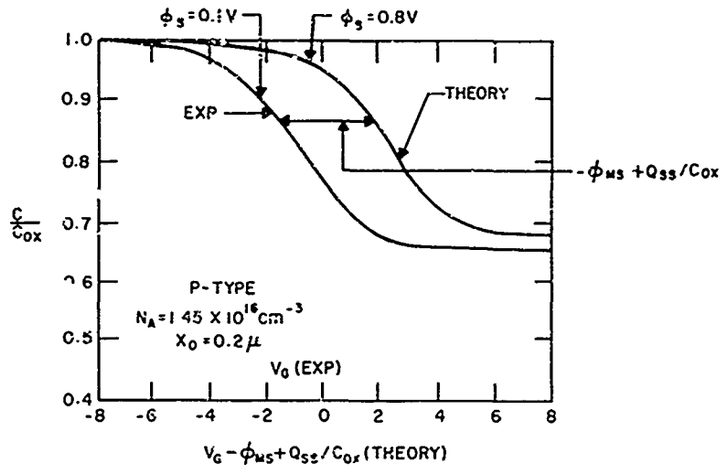


Figure I-8. Comparison of Theoretical and Experimental Curves of C/C_{ox} versus Gate Voltage for a P-Type Si MOS Capacitor

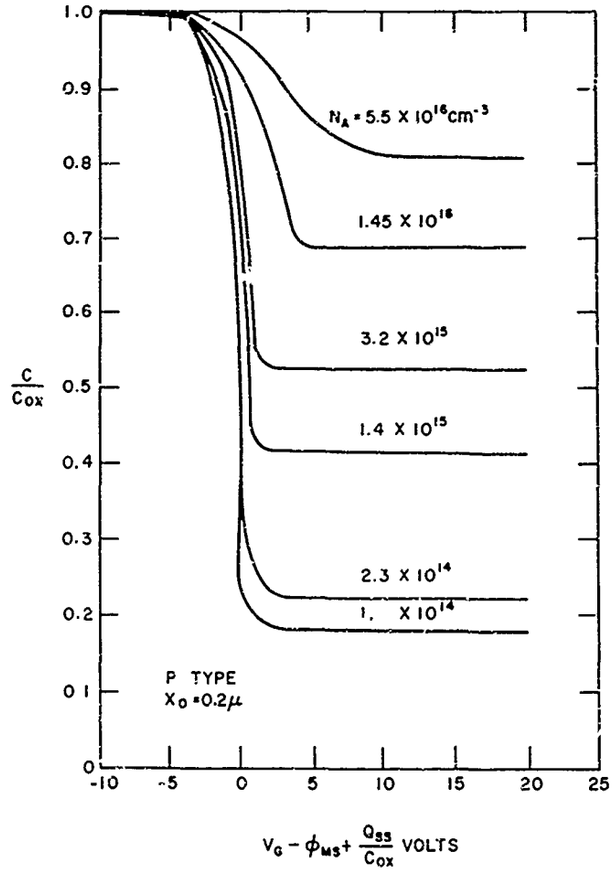


Figure I-9. Dependence of C on N_A for a P-Type Si MOS Capacitor

for capacitors with acceptor concentration from 1.1×10^{14} to 5.5×10^{16} cm^{-3}. The inversion mode capacity can be used to determine N_A in the Si surface layer.

5. DENSITY OF STATES AT THE SiO_2-Si INTERFACE

The measurement of MOS capacitance as a function of applied voltage will yield only a limited amount of information about the charge states in the oxide at the SiO_2-Si interface.[7,8] The interpretation of the experimental results is, in most cases, ambiguous. To illustrate the difficulties encountered, consider the case of a monoenergetic charge site uniformly distributed throughout the oxide, as shown in Figure I-10a. As seen from the Si, the energy of a particular trap site with respect to the bottom of the Si conduction band will obviously be a function of the electric field in the oxide, E_{ox}, and the distance of the site from the interface. The sites within X_m (~20Å) of the interface are filled and thus for a field $\sim 10^6$ V cm^{-1} the energy range ΔE over which the "monoenergetic" charge level is spread is $e\Delta EX_m \sim 0.2$eV (Figure I-10b). Thus a monoenergetic trap level in the oxide yields an effective density of surface states. If a monoenergetic trap level were the only level present in the oxide, the experimental results could be interpreted relatively straightforwardly. However, in practice, several levels at least are likely to be present, and the resolution of these levels from the information obtained from

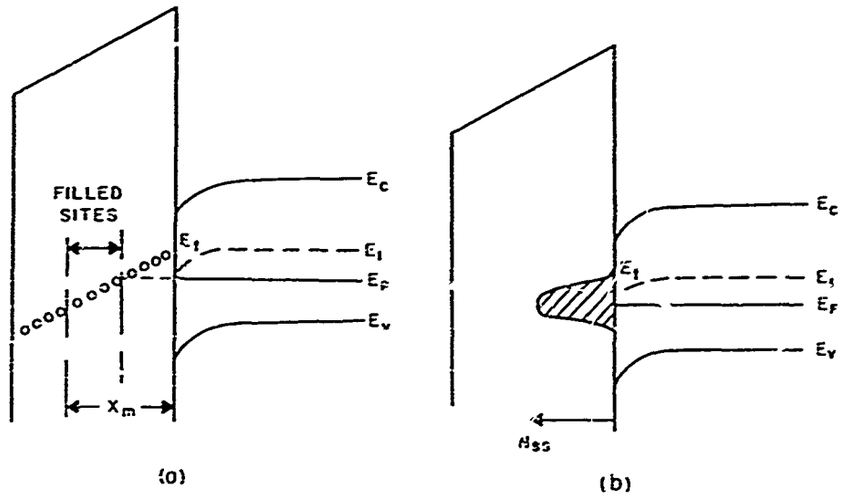


Figure I-10. Monoenergetic Trap Level Uniformly Distributed in SiO_2
(a) Energy diagram
(b) Effective density of states in the oxide as seen from the Si

capacity measurements alone is virtually impossible. Only an effective density of states can be determined.

The procedure for calculating the effective density of states from capacity variation with applied potential can be obtained from a slight extension of the MOS capacitance theory outlined above. The procedure is lengthy, cumbersome, and of limited accuracy and requires the use of a computer if many devices are to be analyzed. It will be assumed that the charge in the oxide surface states, Q_{ss}, is able to follow slow (d-c) variations in the surface potential, φ_s, caused by variations in the applied gate voltage, V_G, but that Q_{ss} cannot follow any a-c changes in φ_s. The density of states, N_{ss}, (per cm^2 per V) is defined by the relation

$$Q_{ss} = q \int N_{ss}\, d\varphi_s$$

where the integration is over the occupied sites (it is assumed that all states below the Fermi level are occupied, all states above are unoccupied). N_{ss} is then given by

$$N_{ss} = \frac{1}{q}\frac{dQ_{ss}}{d\varphi_s} \tag{13}$$

and Q_{ss} must now be found as a function of φ_s.

Equation (10) can be rewritten as

$$C = \frac{C_{ox}\left(dQ_s/d\varphi_s\right)}{\left(dQ_s/d\varphi_s\right) - C_{ox}} \tag{14}$$

where it has been assumed $(dQ_{ss}/d\varphi_s) = 0$ at the a-c frequency used for measurement.

Hence for each bias voltage, V_G, the corresponding value of C can be used with the known value of C_{ox} to determine $dQ_s/d\varphi_s$. From equation (2)

$$\frac{dQ_s}{d\varphi_s} = -\frac{q}{kT}\frac{\varphi_s}{|\varphi_s|}\frac{L_D n_i \sqrt{2}\left[\sinh q\left(\varphi_s - \varphi_F\right)/kT + \sinh q\varphi_F/kT\right]}{\left[\cosh q\left(\varphi_s - \varphi_F\right)/kT + \cosh q\varphi_F/kT + \left(q\varphi_s/kT\right)\sinh q\varphi_F/kT\right]^{1/2}} \tag{15}$$

If equation (15) is used to plot $dQ_s/d\varphi_s$ as a function of φ_s for the particular MOS capacitor, then, from the graph, φ_s can be determined for the values of $dQ_s/d\varphi_s$ calculated from equation (14). The φ_s values then obtained can be used with equation (2) to determine the corresponding values of Q_s, yielding, for a particular MOS capacitor, φ_s, Q_s, C_{ox}, and φ_{Ms} values for each V_G value. Equation (9) can then be used to find the corresponding values of Q_{ss}. A plot of Q_{ss} versus φ_s can

then be used to determine $dQ_{ss}/d\varphi_s$ graphically to give N_{ss} as a function of φ_s through equation (13).

6. USE OF MOS EQUIVALENT PARALLEL CONDUCTANCE TO DETERMINE THE DENSITY OF FAST SURFACE STATES

Nicollian and Goetzberger[5] have demonstrated a method for determining the density of fast surface states at the SiO_2-Si interface. They have shown that the equivalent parallel conductance of a MOS capacitor arises almost entirely from these fast surface states with virtually no contribution from the oxide or depletion layers. Under these circumstances the MOS equivalent circuit shown in Figure I-11 is applicable for frequencies above 100 Hz for which the minority carriers cannot follow the a-c signal. C_{ox} is again the oxide capacitance, C_D is the capacitance of the depletion layer, and C_{ss} and R_{ss} are the capacitance and resistance associated with the fast surface states.

The measured admittance of the MOS capacitor Y_o may be written as

$$Y_o = G_o + j\omega C_o$$

where G_o is the measured conductance, C_o the measured capacitance, and ω the angular frequency of the measured signal. Converting Y_o to an impedance, Z_o, results in

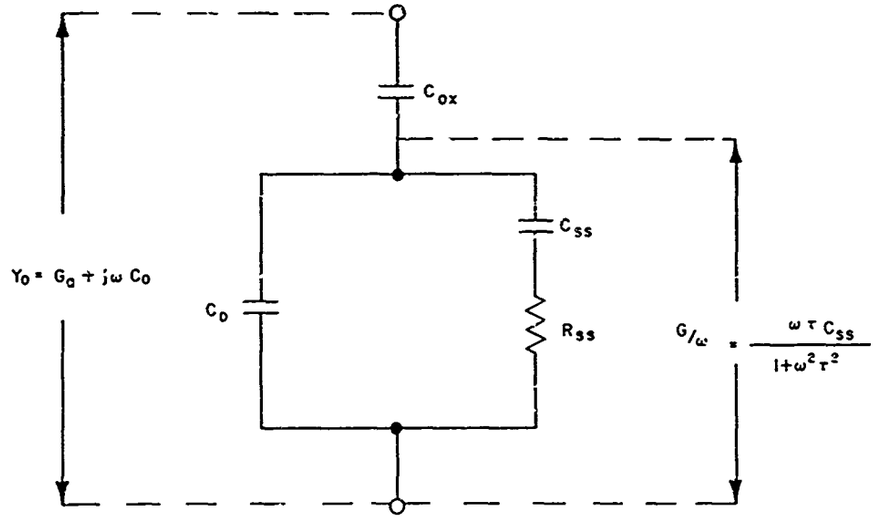


Figure I-11. MOS Equivalent Circuit Including Effect of Loss Due to Fast Surface States

$$Z_o = \frac{1}{Y_o} = \frac{G_o}{G_o^2 + \omega^2 C_o^2} - \frac{j\omega C_o}{G_o^2 + \omega^2 C_o^2}$$

Subtracting the impedance of the oxide from Z_o gives the impedance, Z, of the parallel combination of C_{SS}, R_{SS}, and C_D,

$$Z = \frac{G_o}{G_o^2 + \omega^2 C_o^2} + j\left(\frac{1}{\omega C_{ox}} - \frac{\omega C_o}{G_o^2 + \omega^2 C_o^2}\right)$$

After converting Z to an admittance, the conductance, G, is given by

$$G = \frac{\omega^2 C_{ox}^2 G_o \left(G_o^2 + \omega^2 C_o^2\right)}{\omega^2 G_o^2 C_{ox}^2 + \left[G_o^2 + \omega^2 C_o (C_o - C_{ox})\right]^2} \tag{16}$$

From the equivalent circuit in Figure I-11, however, G is given by

$$\frac{G}{\omega} = \frac{C_{SS}\omega\tau}{1 + \omega^2\tau^2} \tag{17}$$

where τ is the time constant of the fast surface states, $R_{SS} C_{SS}$. Equation (17) is very useful since it does not involve C_D. The quantity G/ω reaches a maximum of $C_{SS}/2$ when $\omega\tau = 1$. Hence, if G is calculated from equation (16) at a particular value of V_G for a series of frequencies and G/ω plotted against ω, then τ and C_{SS} may be determined directly. The density of states, N_{SS}, can be found from equation (13) since $C_{SS} = dQ_{SS}/d\varphi_S$. The process is repeated at other values of V_G (i.e., φ_S) to yield N_{SS} as a function of φ_S.

7. ION TRANSPORT IN SiO_2 FILMS

When devices employing a SiO_2 film are heated or irradiated under operating conditions a positive charge is often observed to collect in the SiO_2 at the SiO_2-Si interface. This charge influences the adjacent Si surface and leads to device degradation. It has been suggested that mobile positive ions in the SiO_2 are responsible. Snow et al.[10] have investigated MOS capacitors purposely contaminated during fabrication with Na^+ or Li^+ ions at the gate-SiO_2 interface. They found that when a capacitor was heated to ~150°C and the gate made positive with respect to the Si, positive charge accumulated at the SiO_2-Si interface. If the capacitor was subsequently shorted or the polarity of the gate potential reversed the positive charge at the SiO_2-S_i interface was reduced to its original value (but never became less than the original value). Snow interpreted these observations as the result of Na^+ or Li^+

transport across the SiO_2 film and developed a model which agrees with their findings.

It can be shown that for an arbitrary charge distribution, $\sigma(x)$, in oxide layer the capacity versus voltage curve for the MOS capacitor will be shifted parallel to the voltage axis by an amount ΔV where

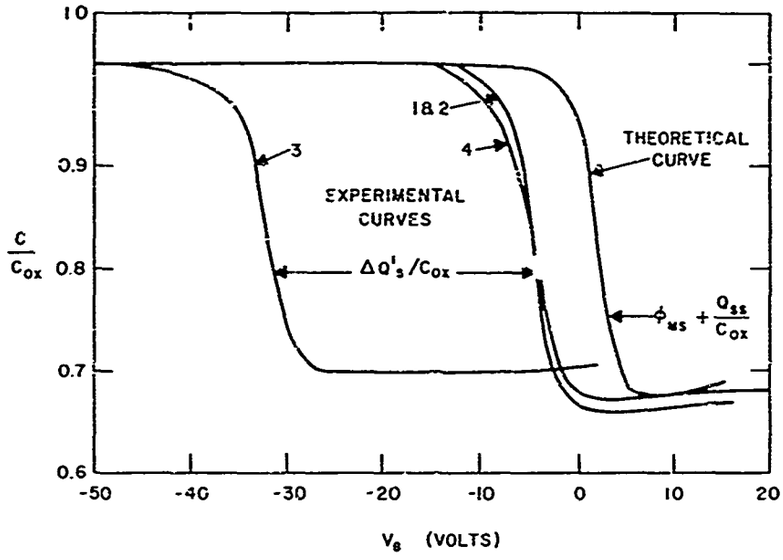
$$\Delta V = \frac{1}{x_o C_{ox}} \int_0^{x_o} x\, \sigma(x) dx$$

The origin of the x coordinates is taken at the gate-SiO_2 interface. Initially all the charge will be located at the gate-SiO_2 interface, $x = 0$, and hence $\Delta V = 0$. As ions move across the oxide ΔV increases and becomes a maximum when all the charge reaches the SiO_2-Si interface, $x = x_o$. The experimental curves obtained by Snow on N_a^+ contaminated MOS capacitors are shown in Figure I-12a and the charge distributions corresponding to the various curves are shown in Figure I-12b.

Curve 1 shows the initial measurement before a bias was applied to the capacitor. The corresponding charge distribution is shown in Figure I-12b(i). After 5 minutes at $V_G = -10V$, $T = 150°C$ the devices were cooled to room temperature and curve 2 observed. It is evident that virtually no change has occurred in the charge distribution, i.e., Figure I-12b(i) applies to curve 2. Curve 3 was obtained after 5 minutes at $V_G = +10V$, $T = 150°C$. Positive charge has apparently collected at the SiO_2-S_i interface as shown in Figure I-12b(ii). After 5 minutes at $T = 150°C$ with the capacitor shorted, curve 4 was obtained indicating the positive charge had largely returned to the gate-SiO_2 interface as shown in Figure I-12b(iii).

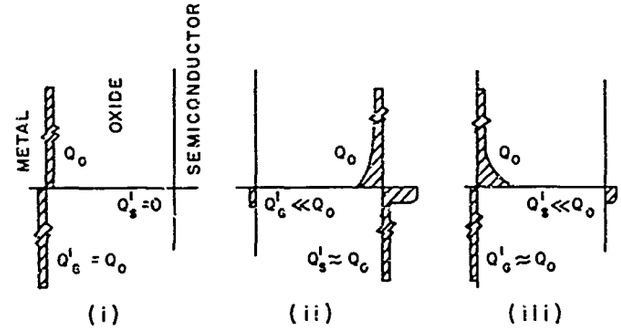
To obtain quantitative expressions for the transport of ions through the SiO_2, Snow divided the oxide layer into two regions, a thin boundary layer at the gate-SiO_2 interface and a bulk region comprising the remainder of the SiO_2. To obtain a model suitable for mathematical analysis the electric field in the boundary layer is assumed to be zero (in practice the average value ~0) and ion motion is due solely to diffusion. In the bulk region the field is taken as constant and ion motion is due to the field and diffusion. According to this model when an ion diffuses out of the narrow boundary region it is rapidly swept to the SiO_2-Si interface, i.e., ions spend most of their time at one or other of the interfaces.

Further analysis of the model yields the time dependence of the buildup of charge at the SiO_2-Si interface, Q'_S, as



(a)

At 150°C: (1) original curve; (2) after 5 minutes at V_G = -10V; (3) after 5 minutes at V_G = +10V; (4) after 5 minutes, shorted



(b)

Charge distribution corresponding to curves in (a): (i) corresponds to curves (1) and (2); (ii) corresponds to curve (3); (iii) corresponds to curve (4)

Figure I-12. D-c Bias Applied to Na-Contaminated MOS Capacitor

$$Q'_s \sim \left(\frac{t}{\tau}\right)^2 \qquad t << \tau$$

$$Q'_s \sim 1 - \frac{8}{\pi^2} \exp\left(-\frac{t}{\tau}\right) \qquad t >> \tau$$

where τ is a characteristic time associated with the oxide ($\tau \sim$ minutes). The expressions above for Q'_s agree quite well with observed time dependence of Q'_s.

8. USE OF MOS CAPACITORS IN THE STUDY OF RADIATION EFFECTS ON SEMICONDUCTOR DEVICES

Semiconductor devices employing a SiO_2 layer degrade when exposed to radiation due to a buildup of positive charge at the SiO_2-Si interface. The charge corresponds to Q_{ss} in a MOS capacitor. Since changes in Q_{ss} may be determined with relative ease from capacity measurements MOS capacitors will be useful in the study of the role played by radiation, in conjunction with electric field and temperature in the buildup of Q_{ss}.

In particular. it is desirable to know the relationships between Q_{ss} and temperature, radiation dose, dose rate and bias voltage as well as any conditions which lead to a decrease (recovery) of Q_{ss}. A knowledge of these relationships may provide a clue to the mechanism by which Q_{ss} accumulates.

If the increase of Q_{ss} is due to the motion of positive ions such as N_a^+, as proposed by Snow, then one might expect the charge to be contained in a set of monoenergic traps near the SiO_2-Si interface. A determination of the effective density of states N_{ss} as described in paragraph 5 would indicate the presence of a monoenergetic trap level. However, if a more complicated set of levels were present, the effective density of states could not be unambiguously interpreted without additional information (such information may be obtainable from optical studies).

MOS-FET's are very sensitive to conditions at the SiO_2-Si interface and hence are very sensitive to radiation. At relatively low radiation doses the turn-on voltages and transconductance of these devices may be seriously degraded. By a fortunate coincidence MOS-FET's contain a built-in MOS capacitor and are thus able to serve the dual purpose of MOS capacitor and device. With the aid of MOS-FET's it should be possible to obtain in the same structure a direct correlation between Q_{ss} and degradation. Such a correlation will be very useful since, strictly speaking, it is the charge Q_{ss} and not the radiation which actually causes degradation.

REFERENCES

1. L. M. Terman, "An Investigation of Surface States at a Silicon/Silicon Oxide Interface Employing Metal-Oxide-Silicon Diodes," Solid-State Electron., Vol. 5, No. 4 (1962) 285.

2. K. Lehovec, A. Slobodskoy, and J. L. Sprague, "Field Effect-Capacitance Analysis of Surface States on Silicon," Phys. Status Solidi, Vol. 3 (1963) 447.

3. A. S. Grove, E. H. Snow, B. E. Deal, and C. T. Sah, "Simple Physical Model for the Space-Charge Capacitance of Metal-Oxide-Semiconductor Structures," J. Appl. Phys., Vol. 35 (1964) 2458.

4. A. S. Grove, B. E. Deal, E. H. Snow, and C. T. Sah, "Investigation of Thermally Oxidised Silicon Surfaces Using Metal-Oxide-Semiconductor Structures," Solid-State Electronics, Vol. 8 (February 1965) 145.

5. E. Deal, A. S. Grove, E. H. Snow, and C. T. Sah, "Recent Advances in the Understanding of the Metal-Oxide-Silicon System," Trans. Metall. Soc. AIME, Vol. 233, No. 3 (March 1965) 524.

6. R. Lindner, "Semiconductor Surface Varactor," B.S.T.J., Vol. 41 (May 1962) 803.

7. F. P. Heiman and G. Warfield, "The Effects of Oxide Traps on the MOS Capacitance," IEEE Trans. on Electron Devices, Vol. ED-12 (1965) 167.

8. K. H. Zaininger and G. Warfield, "Limitations of the MOS Capacitance Method for the Determination of Semiconductor Surface Properties," IEEE Trans. on Electron Devices, Vol. ED-12 (1965) 179.

9. E. H. Nicollian and A. Goetzberger (to be published).

10. E. H. Snow, A. S. Grove, B. E. Deal, and C. T. Sah, "Ion Transport Phenomena in Insulating Films," J. Appl. Phys., Vol. 36 (1965) 1664.

ATTACHMENT II

THE EFFECTS OF FAST-NEUTRON IRRADIATION ON THE SATURATION VOLTAGE OF SILICON PLANAR EPITAXIAL TRANSISTORS*

1. EFFECTS OF FAST-NEUTRON IRRADIATION ON SILICON

The study of the effects of fast-particle bombardment on semiconductors had its historical beginning in 1947 with the discovery that the exposure of germanium and silicon specimens to reactor irradiations and cyclotron particles affected their conductivity and Hall coefficient.[1] This discovery was made concurrently at Purdue University by K. Lark-Horovitz and at Oak Ridge by W. E. Johnson. Since the changes in conductivity and Hall coefficient were too great to be attributed to impurities introduced by nuclear reactions, it was concluded that the observed change in both concentration and type of charge carrier was caused by localized energy states associated with radiation-induced lattice disorder. This view was later confirmed when it was demonstrated that most of the changes could be removed by thermal annealing near 500°C.[2]

Since this early work of Johnson and Lark-Horovitz, considerable effort has been devoted to the study of radiation effects in a variety of semiconductors of diamond type lattice using a number of radiation sources. There are several reasons for the high degree of interest in this field. The physicist sees in radiation defects the opportunity to study the simplest possible defects in crystalline solids. For example, radiation has been used in the determination of the displacement energy, therefore justifying the assumption that lattice atoms sit in a simple potential well of depth α.[1] Radiation can be used as a valuable tool for altering the properties of semiconducting materials to produce new types of semiconducting materials with predictable properties. As will be discussed later, the interaction of irradiation with semiconducting materials can alter such properties of the material as charge carrier concentration and type, mobility, minority-carrier lifetime, minority-carrier trapping processes, and other associated properties.[1]

Another increasingly important reason for studying the effects of irradiation on semiconductor material is the increasing application of semiconductors in environments that are subject to radiation fields. Today there is a need for semiconductor devices that will be able to operate reliably for long periods of time in the

---

*This attachment is the thesis submitted by E. A. Overstreet in partial fulfillment of the requirements for the degree of Master of Science in the Department of Engineering, Graduate School of Arts and Sciences, Duke University, Durham, N.C.

presence of radiation. Such devices are required for use in military applications and in communication satellites which must operate in the Van Allen belts.

Semiconductors have proven to be much more sensitive to radiation than metals.[3] The reason is that specimens of semiconductor materials can be prepared with a lower concentration of carriers in the conduction band, and consequently the introduction of a few traps or donors by bombardment with energetic particles can lead to relatively large changes in the carrier concentration. As a result of this property, semiconductors are used as particle detectors.[4]

The basic neutron radiation damage process results from the collision of an incident high-energy particle with an atom in the semiconductor crystal. Since the neutron is an uncharged particle, the initial interaction will be an elastic collision with the nuclei of the lattice atom, rather than a Coulomb interaction characteristic of charged incident particles. If, in the elastic collision, the lattice atom receives an energy in excess of the displacement energy E_d, the binding energy of the normal lattice site may be overcome and the atom will be displaced from its normal lattice site. Therefore, a lattice vacancy is created and, since in most circumstances the recoiling atom comes to rest in a nonequilibrium or interstitial state, an interstitial atom is created. Thus the primary process involved in fast-neutron bombardment is production of Frenkel defects (pairs of vacant lattice sites and interstitial atoms).

As a secondary process, the displaced or recoil atom may receive sufficient energy from the primary collision for it to create vacant lattice sites by subsequent collisions with other lattice atoms. This process can continue until the energy of each particle has been degraded to the point where E_d cannot be transferred in subsequent collisions. This displacement cascade accounts for most of the defects produced during bombardment of semiconductors with fast neutrons. It should be noted that although the primary collision of the neutron and the lattice atom does not involve Coulomb interaction, the recoiling atoms usually become ionized and Coulomb interactions dominate in secondary and subsequent displacements.

In addition to the displacement of atoms from their normal lattice sites, fast-neutron irradiation can produce other effects such as thermal spikes and nuclear transmutations.[3] A thermal spike results when a lattice atom is struck by a fast particle but does not receive enough energy to be removed from its normal lattice site. Instead it will vibrate with large amplitude and will rapidly transfer energy to its neighbors, thus creating a region of high temperature (~1000°K). The region will rapidly expand with a drastic reduction in temperature. The duration of the thermal spike is only 10^{-10} to 10^{-11} seconds. Nuclear transmutations are equivalent to introducing impurity atoms in the conventional manner, and occur to an appreciable extent only in materials of high cross sections. Since the absorption cross section of semiconductors is very small[5] (1×10^{-25} cm^2 for Si), transmutation

effects may be neglected. Therefore, permanent changes in the semiconductor properties with irradiation can be attributed almost entirely to radiation-induced lattice disorder.

The maximum kinetic energy that can be transmitted to a lattice atom by an incident particle occurs in a head-on collision. In the case of massive particle irradiation*

$$E_m = \frac{4 M_1 M_2}{\left(M_1 + M_2\right)^2} E \qquad (1)$$

where E_m is the maximum energy transferred by a moving particle of mass M_1 and energy E to a stationary atom of mass M_2. In order for an atom to be displaced from its normal lattice site, $E_m \geq E_d$. Therefore

$$E_t = \frac{\left(M_1 + M_2\right)^2}{4 M_1 M_2} E_d \qquad (2)$$

where E_t is the minimum or threshold energy that the incident particle must possess in order to displace an atom from its normal lattice site.

A value of 25 ev is usually assumed as a reasonable value for E_d for most materials.[6] However, early experiments with germanium and silicon have led to a value of displacement energy of 31 ev.[7] Recent experiments have shown that damage could be produced well below this supposed value of 31 ev, but with decreasing probability.[4] Rather than define the displacement energy E_d as a value such that all atoms are displaced which receive energy greater than E_d and no atoms are displaced which receive energy less than E_d, it would be more accurate to define an energy-dependent displacement probability similar to the curve shown in Figure II-1.

As mentioned previously, the observed change in carrier concentration of semiconductors with irradiation was caused by localized energy states associated with radiation-induced lattice disorder. James and Lark-Horovitz[8] proposed a model in which the interstitial-vacancy pair provided two donor levels, attributed to the first and second ionization potentials of the interstitial, and two acceptor states whose levels are taken to be the energies required to place one and two electrons respectively into a vacant site. With increasing amounts of irradiation, the induced donors and acceptors will force the Fermi level toward a limiting value determined by their positions in the forbidden-energy gap. When silicon is irradiated, its conductivity approaches the intrinsic value, independent of whether the silicon was initially N- or P-type.[1] The expression for resistivity (ρ) is:

*See Appendix E for List of Symbols.

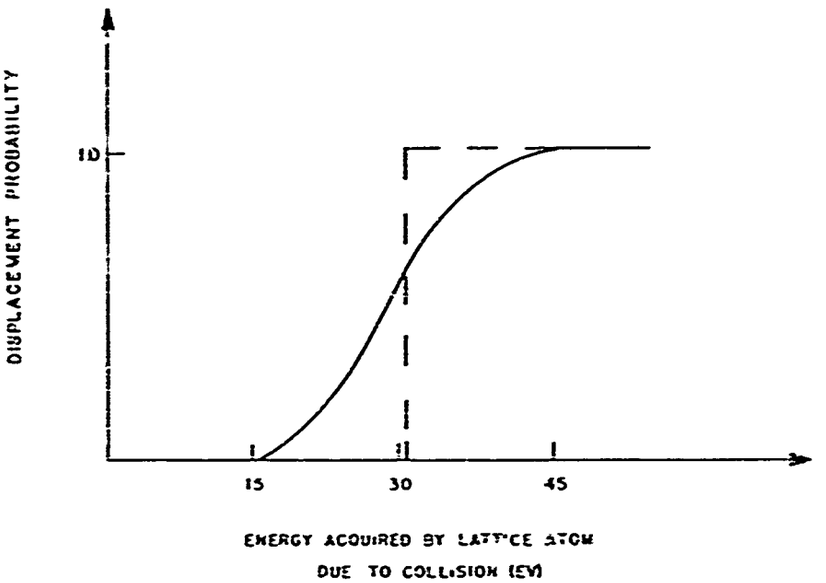


Figure II-1. Displacement Probability

$$\rho = \frac{1}{q\left(\mu_n n + \mu_p p\right)} \text{ ohm-cm} \tag{3}$$

where

μ_n, μ_p = mobility of electrons and holes
n,p = electron and hole concentration
q = electronic charge

The increase in the resistivity of silicon toward intrinsic resistivity can be attributed primarily to the reduction in the majority-carrier concentration by the interstitial atoms and vacancies. A secondary contributor to the increase in resistivity is a reduction in carrier mobility due to the increase of charged scattering centers produced by the ionized recoil atoms.

Since the resistivity of N- or P-type silicon approaches the intrinsic value with irradiation, it was suggested that the two donor levels and two acceptor levels are distributed symmetrically about the center of the forbidden band[1] as shown in Figure II-2.

The actual positioning of the acceptor and donor levels in irradiated silicon has not been successfully established at the present time. However, calculations of carrier removal rates by assuming a donor level at ~0.25 ev below the conduction

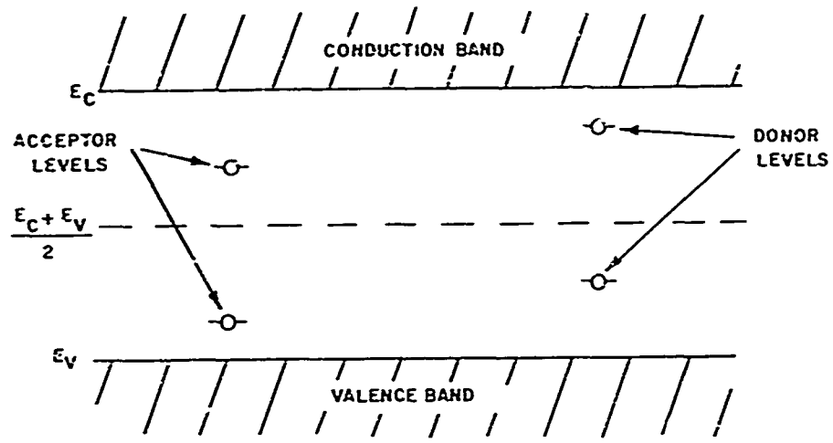


Figure II-2. Energy Level Diagram for Irradiated Silicon Based on James and Lark-Horovitz Model

band and an acceptor level at ~0.055 ev above the valence band give results that compare satisfactorily with experimental measurements.[4] It is seen that if the levels are introduced deep within the band, where thermal ionization is inefficient, then N-type silicon will lose its conduction electrons to those acceptor levels or electron traps which lie below the Fermi level, and P-type silicon will lose its holes from the valence band to those donor levels or hole traps which lie above the Fermi level. The Fermi level is shifted toward the center of the forbidden-energy band, i.e., toward the intrinsic position.

It should be noted that radiation-induced lattice defects in other semiconductor materials are not located symmetrically about the center of the forbidden-energy band. For example, studies of radiation effects on germanium showed that N-type germanium was converted to P-type and P-type increased in conductivity with exposure.[1] This indicates that only radiation-induced acceptor levels have an appreciable effect on germanium. The donor levels lie so near the top of the filled band that they have no observable effect on the carrier concentration.[9] The Fermi level approaches a position ~0.2 ev above the valence band at high levels of irradiation.

Recent work has shown that a more complicated model than the simple four-level model of James and Lark-Horovitz is needed to explain all of the effects of fast-neutron irradiation on semiconductors. Clusters of interstitials and especially vacancies may be produced which have different effects from isolated vacancies and interstitials. A multitude of levels, rather than the four predicted by the earlier models, has been discovered in irradiated semiconductors. Nevertheless, the

mechanism by which majority carriers are removed is essentially the same as was explained for the simpler four-level models.

Another important effect of neutron irradiation of semiconductor materials is a reduction in minority-carrier lifetime. The reduction in minority-carrier lifetime is attributed to the introduction of recombination centers and temporary minority-carrier trapping levels in the forbidden-energy band. Based on a single recombination level and on a low injected carrier density compared to the free carrier density, Shockley-Read-Hall[10] have shown that the carrier lifetime (τ) is given by

$$\tau = \frac{\tau_{pr}\left(n + n_r\right) + \tau_{nr}\left(p + p_r\right)}{n + p} \text{ sec} \tag{4}$$

where

n, p = electron and hole concentrations

n_r, p_r = electron and hole concentrations that would exist if the Fermi level coincided with the level of the recombination centers

τ_{nr} = lifetime of electrons in highly P-type material

τ_{pr} = lifetime of holes in highly N-type material.

Both τ_{nr} and τ_{pr} are limited by the number of recombination centers and can be expressed as

$$\tau_{pr} = \frac{1}{N_R C_p} \text{ sec} \tag{5}$$

$$\tau_{nr} = \frac{1}{N_R C_n} \text{ sec} \tag{6}$$

where N_R is the concentration of recombination centers in cm^{-3} and C_p and C_n are the capture probabilities of the centers for holes and electrons respectively in cm^3 per second. Thus Equation (4) can be written as

$$\tau = \frac{\left(1/C_p\right)\left(n + n_r\right) + \left(1/C_n\right)\left(p + p_r\right)}{N_R\left(n + p\right)} \text{ sec} \tag{7}$$

This relation indicates that as the number of recombination centers N_R increases, the carrier lifetime τ decreases. Beck, Paskell, and Peet[11] indicate that the number of recombination centers introduced is directly proportional to the number of incident neutrons.

Another means of expressing the change in minority-carrier lifetime is given by[12]

$$\frac{1}{\tau_F} = \frac{1}{\tau_i} + \frac{\phi}{K} \text{ sec}^{-1} \tag{8}$$

where

ϕ = integrated fast-neutron flux (neutrons/cm^2)
τ_i = lifetime before irradiation (sec)
τ_F = lifetime after irradiation (sec)
K = radiation damage constant, which for P-type silicon is $3.2 \pm 1.1 \times 10^6$ and for N-type silicon is $2.8 \pm 0.8 \times 10^6$ (nvt-sec)

The damage constant is determined by observing the decay of τ with irradiation in the semiconductor material. The damage constant is dependent on the total neutron dosage and on the properties of the particular semiconductor, such as energy gaps and doping levels.

The effects of fast-neutron bombardment on semiconductor materials can be extended to semiconductor devices, in particular, transistors. Neutron irradiation causes changes in three parameters which affect transistor operation, i.e., resistivity ρ, minority-carrier lifetime τ, and surface recombination velocity S.[13] For neutron flux levels of 10^{14} neutrons/cm^2 or less, the reduction in minority-carrier lifetime is the predominant cause of permanent damage in a transistor.

2. THE SILICON PLANAR EPITAXIAL TRANSISTOR

The silicon planar expitaxial transistor has been found to possess highly desirable properties for use as a general switching device.[14] Its advantages as a switching device can be seen by examining its manufacturing process.

First the collector is formed by growing a very thin, lightly doped, single-crystal layer of semiconductor material on a very heavily doped crystal layer of the same type.[15] This thin, lightly doped layer is known as the epitaxial layer and has a much higher resistivity than the substrate on which it is formed. The epitaxial layer is then covered with an oxide coating. The oxide coating is removed in controlled areas by a photoetching process. The base is then diffused to the collector in the area where the oxide coating has been removed. This method of accurately controlling the diffusion area is known as the planar process. The emitter is then diffused into the base by repeating the planar process used in the base diffusion. Metal contacts are made to the collector, base, and emitter, and the surface of the semiconductor wafer is covered with an oxide coating. The resulting structure is shown in Figure II-3.

In commercial power transistors the emitter is not usually a simple configuration as shown in Figure II-3. Rather, it is usually composed of several strips interleaved with base strips or of a star or grid geometry in order to increase the

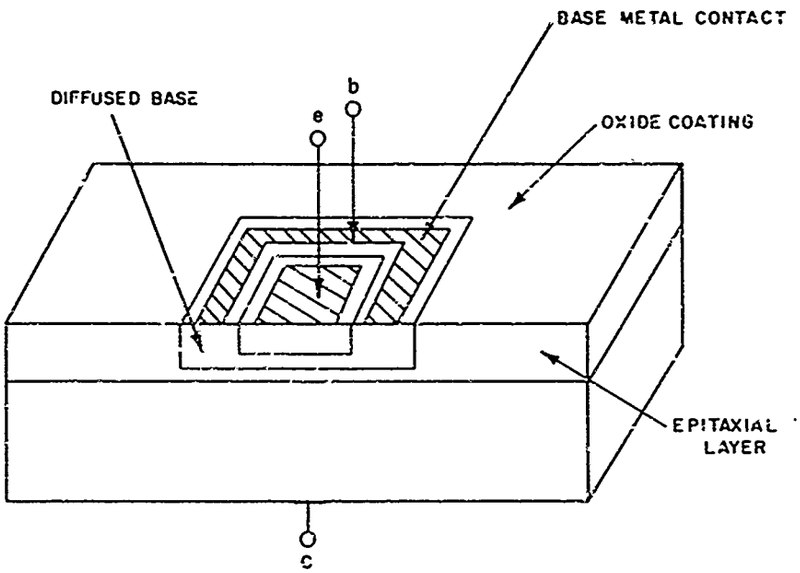


Figure II-3. N-P-N Planar Epitaxial Transistor

ratio of emitter periphery to emitter area to reduce the current crowding effect. This will be discussed later.

Now consider the advantages of the planar epitaxial transistor for switching applications. Since the base is diffused into the collector, the base width can be made smaller (order of microns) than the base width obtained by other junction-forming methods. The narrow base gives a lower series resistance and a smaller base-transit time, resulting in improved frequency response and lower saturation voltage. In switching applications, the higher f_T and low $V_{CE(SAT)}$ of the diffused base transistors give a favorable advantage over alloy and other types of transistors.

Before considering the advantages of the epitaxial layer, it will be necessary to consider the dependence of certain parameters of a switching transistor on the collector doping concentration. It is desirable to have a collector of high resistivity to give a high collector-base breakdown voltage BV_{CBO} and a low collector junction capacitance. However, increasing the collector resistivity increases the voltage drop across its ohmic resistance which increases $V_{CE(SAT)}$. In addition, the high minority-carrier lifetime associated with high resistivity establishes a significant amount of stored charge in the collector, resulting in high switching times. Thus, it can be seen that the design of the collector in switching transistors is a compromise between BV_{CBO}, $V_{CE(SAT)}$, C_{TC}, and switching time.

With a thin high-resistivity epitaxial region in the collector, this compromise is not necessary. The high-resistivity epitaxial layer gives a high collector

breakdown voltage and very low capacitance per unit area. Because the epitaxial layer is very thin (typically 15 to 20 microns), the total series resistance affecting $V_{CE(SAT)}$ is small, since the resistance of the heavily doped substrate is negligible. In addition, the collector-stored charge is reduced considerably because the high recombination rate in the heavily doped substrate reduces the minority-carrier diffusion length to the thickness of the epitaxial region, which is several times less than the diffusion length corresponding to the actual lifetime in the epitaxial region.

The effect of the oxide coating on the surface of the semiconductor wafer is to reduce surface recombination and reduce leakage currents.

For the investigation in this thesis, three types of silicon power transistors " th planar epitaxial construction were used. They are Motorola SF2553 and SF2585 and Clevite 3TX002. These transistors were investigated because they were known to have the desired parameters for particular switching applications, but it was not known how their parameters would change with fast-neutron irradiation.

The Motorola SF2553 transistor is a prototype NPN silicon power transistor with a measured base width of ~1.3 μ. The Motorola SF2585 NPN silicon transistor is similar to the SF2553, except that its base width is approximately 1 μ. The Motorola transistors are rated at 5 amperes collector current and they are contained in a TO-3 package. The Clevite 3TX002 is an NPN silicon power transistor with a base width of approximately 1 μ. It is rated at 5 amperes collector current and a maximum power dissipation of 60 watts. The emitters of the Motorola and the Clevite transistors are composed of strips with base strips between the emitter strips. This arrangement increases the emitter-periphery-to-emitter-area ratio. The essential data for determining the switching characteristics of these transistors is given in Table II-1.

3. EXPERIMENTAL TECHNIQUES AND DATA

The fast-neutron irradiation for this investigation was performed at the Sandia Pulse Reactor Facility (SPRF) in Albuquerque, New Mexico. The devices irradiated included four Clevite 3TX002, five Motorola SF2553, and five Motorola SF2585. During irradiation, the devices under study were mounted on a styrofoam block as shown in Figure II-4. The mounting surface was cut on an 8-inch radius in order that all devices would be equally distant from the center of the cylindrical reactor when the styrofoam block was placed at the proper distance from the reactor. The amount of integrated fast-neutron flux received by the devices was controlled by the distance from the reactor at which they were mounted and by the number of radiation pulses which they received.

Table II-1

TRANSISTOR PROPERTIES

	SF2585	SF2553	3TX002
Base width ($\mu = 10^{-4}$ cm)	1	1.3	1
Base resistivity (Ω-cm)	0.01	0.01	0.01
Epitaxial layer thickness (μ)	15	15	20
Epitaxial layer resistivity (Ω-cm)	2	2	5
Emitter-base perimeter (inch)	0.4	0.4	1.0
Emitter thickness (μ)	2	1.7	3.0
Collector body thickness (inch)	0.005	0.005	0.005
Collector body resistivity (Ω-cm)	0.0001	0.0001	0.0001
Emitter resistivity (Ω-cm)	0.0001	0.0001	0.0001
Emitter-base contact spacing (μ)	25	25	30
Average sheet resistance of base region underneath emitter (Ω/sq)	~2000	~1200	~2400
Average sheet resistance of base region not underneath emitter (Ω/sq)	100	100	100

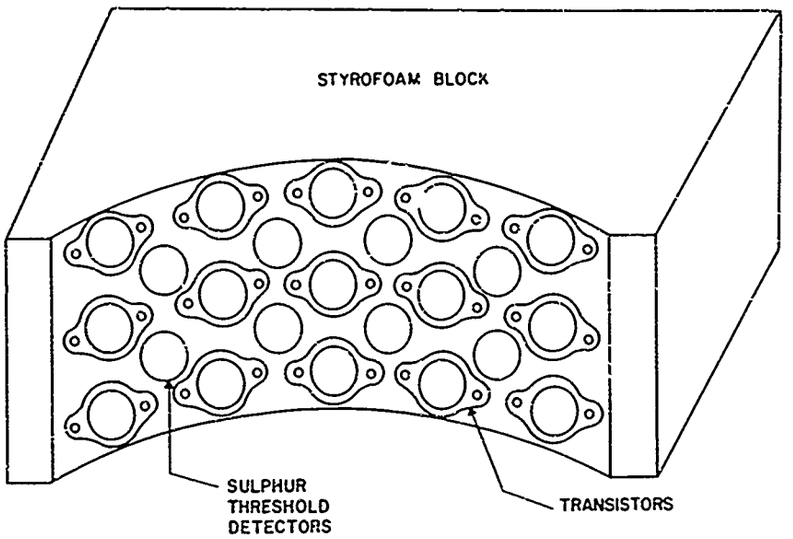


Figure II-4. Mounting of Transistors for Irradiation

Sulfur threshold detectors exposed with the transistors were used to determine the integrated fast-neutron flux (E > 10 Kev). It was discovered that the flux received by the transistors varied slightly with angle over the mounting surface. The data presented take this variation of flux into account. The total integrated fast-neutron flux received by the transistors was approximately 10^{14} neutrons/cm^2.

Measurements were made of common-emitter d-c current gain h_{FE}, collector-emitter saturation voltage $V_{CE(SAT)}$, and base-emitter saturation voltage $V_{BE(SAT)}$ at six different levels of irradiation. All measurements were made using a Baird Atomic Power Transistor Pulse Unit. This instrument applies 300-microsecond pulses to the base of the transistor at a repetition rate of 60 cycles per second. The resulting base and collector current pulses are peak detected and compared to d-c measuring values. In this way the average power dissipation is reduced by a factor of 50. Thus the effect of temperature change on the transistor parameters can be neglected.

Graphs of direct current gain versus exposure for the various transistors are shown in Figures II-5 through II-7, where h_{FE} is given for collector currents of 1 and 3 amperes. Although this thesis is concerned primarily with saturation voltage and not current gain, data on d-c current gain versus exposure are important in the analysis of saturation voltage for two reasons. First, it will be shown in Chapter 4 that saturation voltage is a function of d-c current gain. Second, d-c current gain versus exposure to irradiation gives a good approximation of minority-carrier

lifetime changes with irradiation. This information will be shown to be important in predicting changes in saturation voltage with irradiation. The total reduction of d-c gain with irradiation is summarized for the three types of transistors in Figure II-8.

Graphs of V_{BE} versus exposure for the transistors are shown in Figures II-9 through II-11. As can be seen, V_{BE} does not change significantly with irradiation up to the flux levels considered in this investigation. Graphs of V_{CE} versus exposure for all the transistors at collector currents of 1 and 3 amperes and an I_C/I_B ratio of five are shown in Figures II-12 through II-14. The reason V_{CE} versus exposure data in the above mentioned figures was presented for individual transistors is to show the consistency of the increase in V_{CE} with exposure for the transistors. Also, the comparison of theory and experimental data in a later chapter will be focused on the particular current conditions shown in these figures.

To show the dependence of V_{CE} versus exposure on base current, graphs of V_{CE} versus exposure for different base currents are presented in Figures II-15 through II-17 as an average for transistors of the same type.

4. THEORETICAL CONSIDERATION OF COLLECTOR-EMITTER SATURATION VOLTAGE AND PREDICTED CHANGE WITH IRRADIATION

An ideal switch would be an open circuit when "off" and a short circuit when "on." A transistor used as a switch does not have this ideal characteristic, but rather has a small voltage drop across it when in the "on" condition. In switching applications, this voltage drop, known as the collector-emitter saturation voltage ($V_{CE(SAT)}$), is of prime importance and should be kept as low as possible for two reasons. First, it is superimposed on the signal that is being transmitted through the "on" transistor, thus representing error in the transmitted signal; second, the saturation voltage represents power lost in the transistor.

By definition, saturation is the condition in which both the emitter and collector junctions are forward biased.[16] For a common-emitter circuit configuration, the base current required to hold the transistor in a saturated condition is given by the following equation

$$I_B > \frac{I_C}{\beta_N} \tag{9}$$

where

$$\beta_N = \frac{\alpha_N}{1 - \alpha_N}$$

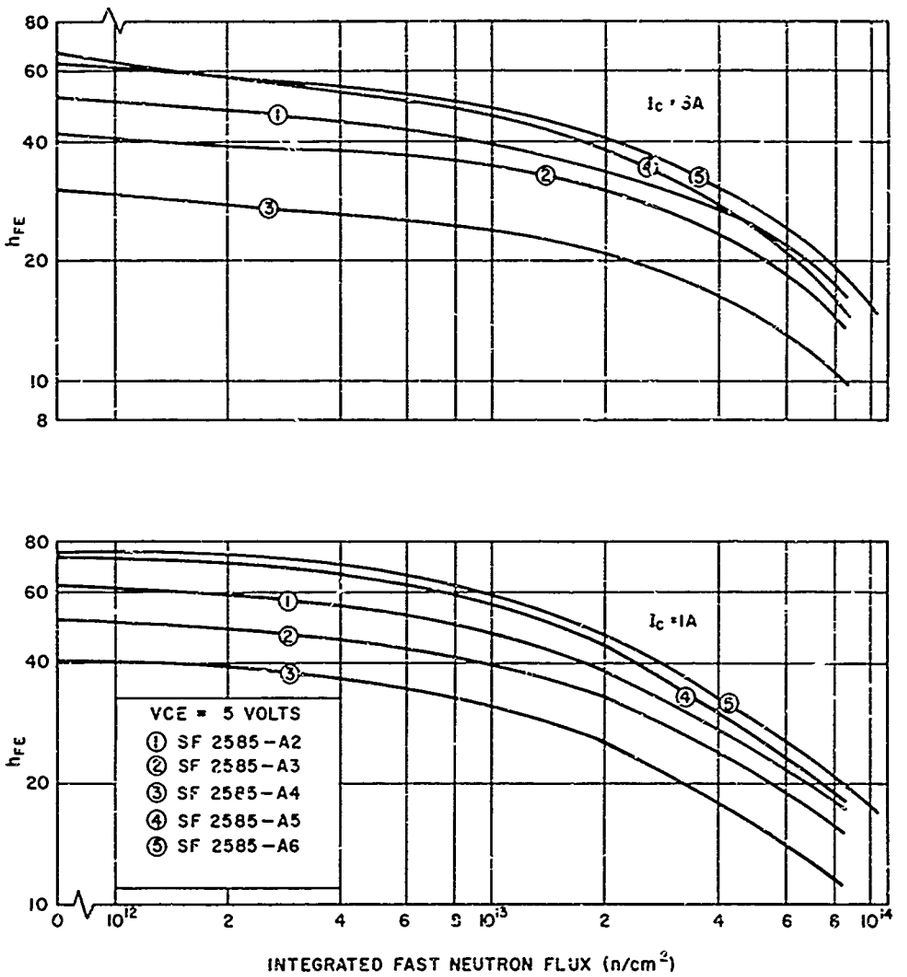


Figure II-5. Decrease in D-c Gain of Motorola SF2585 with Fast-Neutron Irradiation

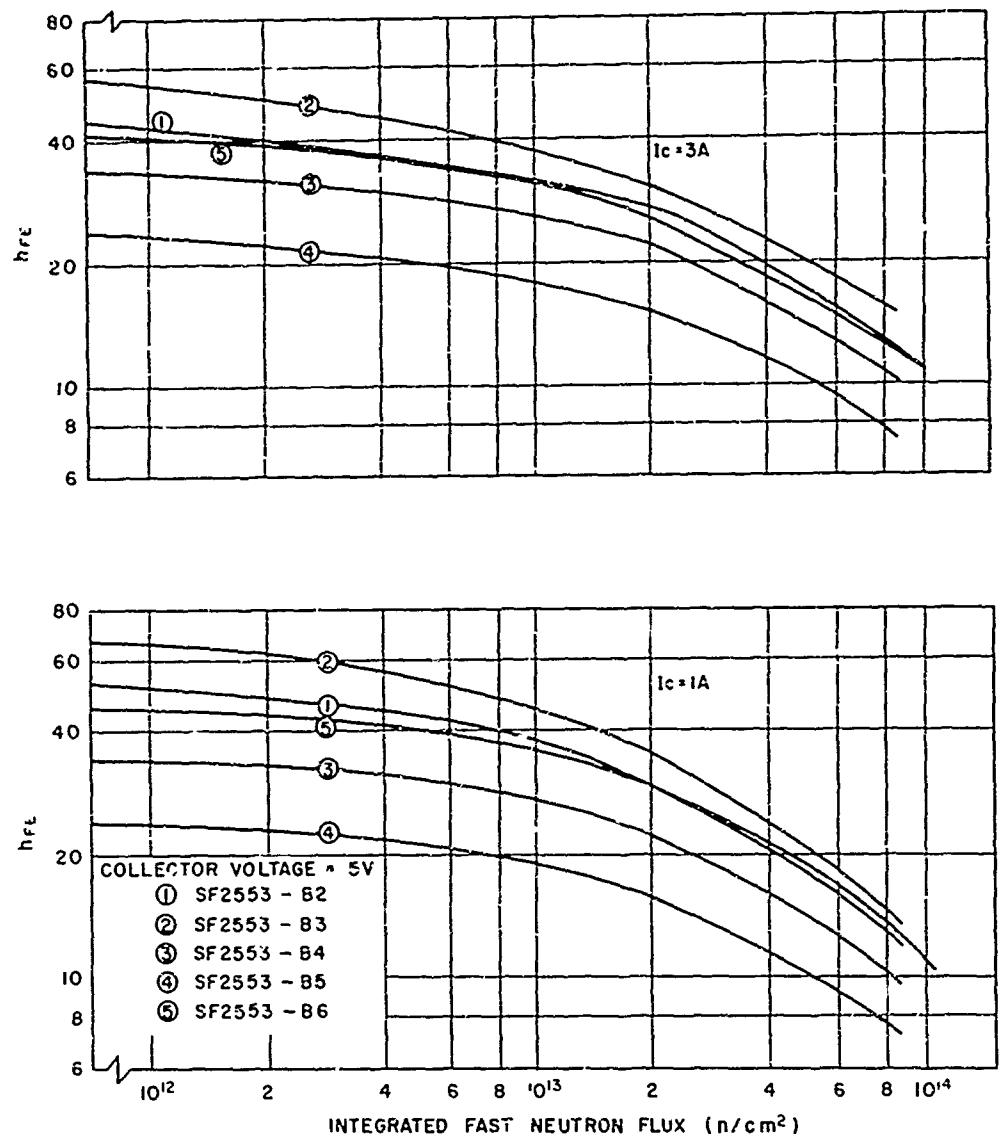


Figure II-6. Decrease in D-c Gain of Motorola SF2553 with Fast-Neutron Irradiation

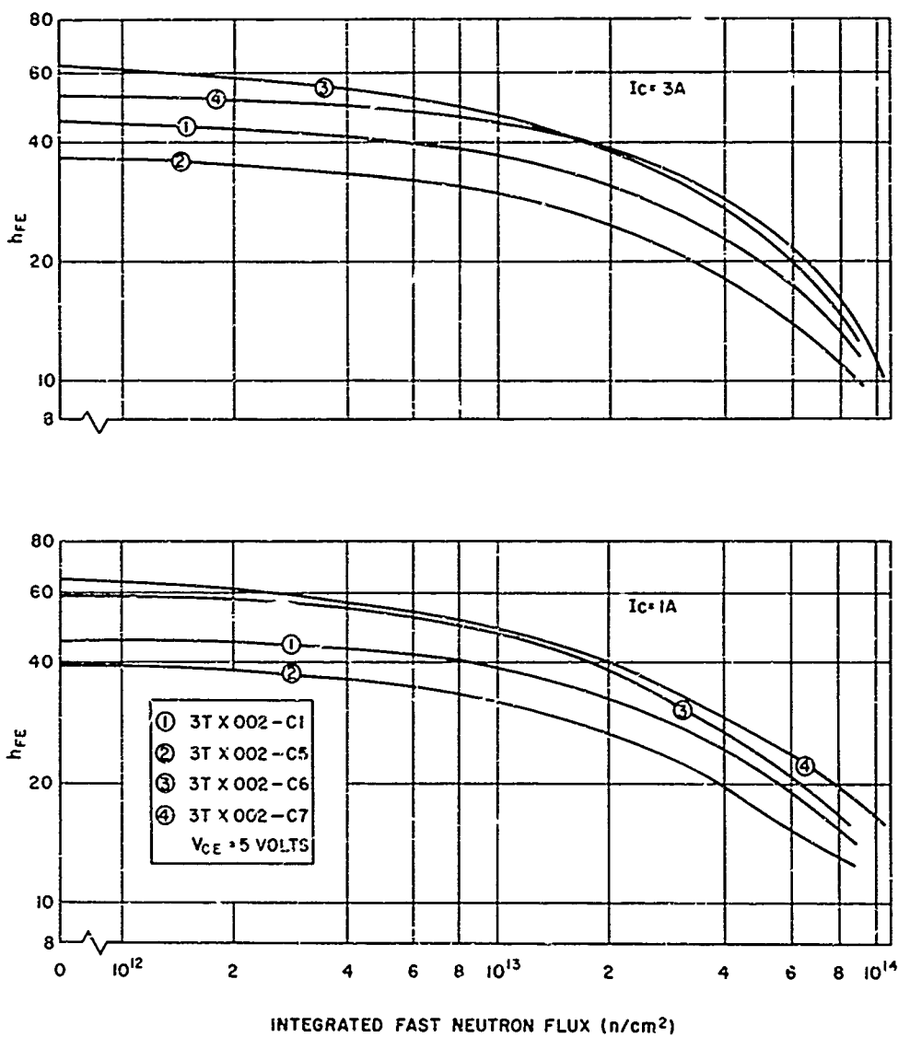


Figure II-7. Decrease in D-c Gain of Clevite 3TX002 with Fast-Neutron Irradiation

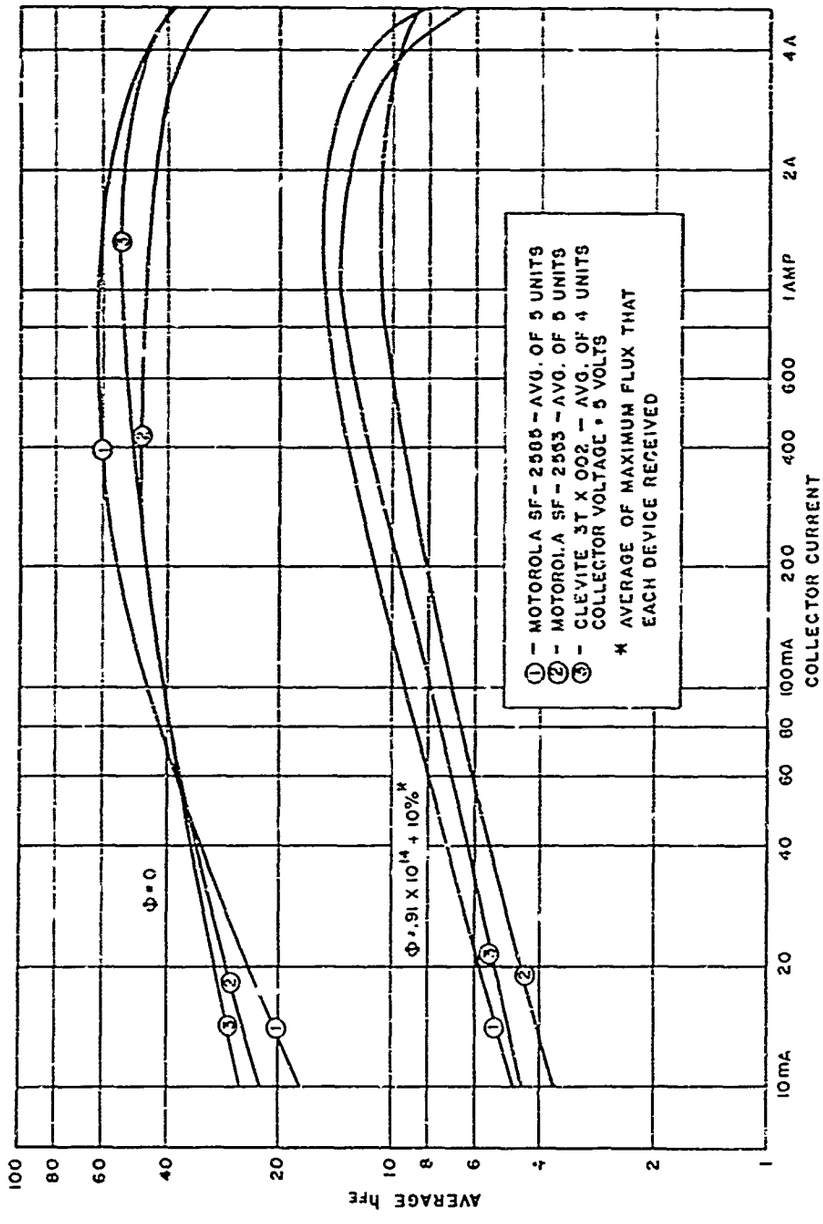


Figure II-8. Average D-c Gain versus Collector Current with Fast-Neutron Exposure as a Parameter

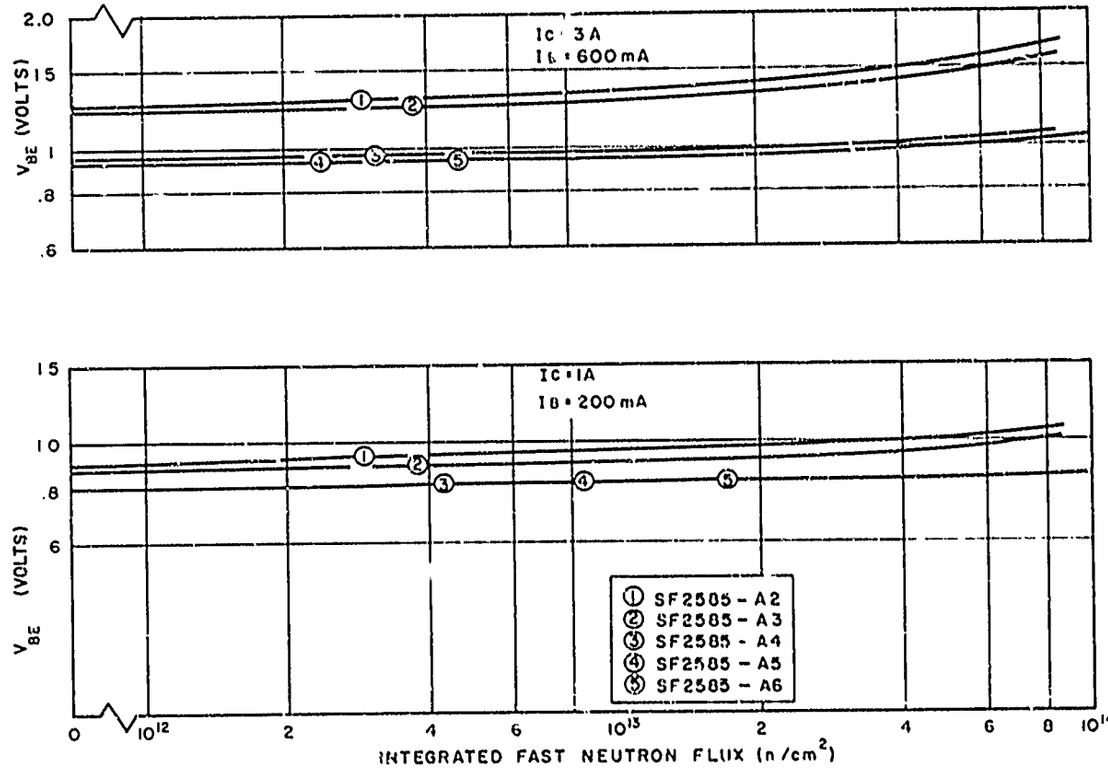


Figure II-9. Change in Emitter-Base Voltage of Motorola SF2585 with Fast-Neutron Irradiation

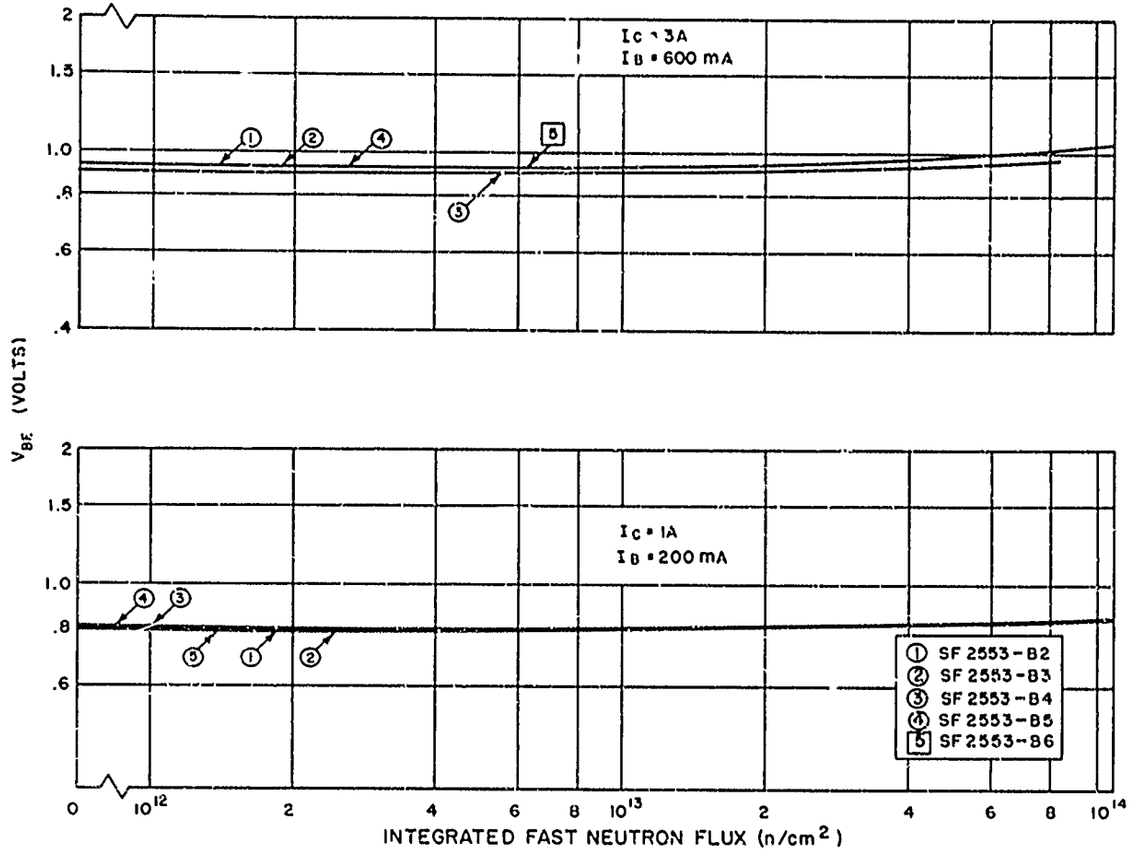


Figure II-10. Change in Emitter-Base Voltage of Motorola SF2553 with Fast-Neutron Irradiation

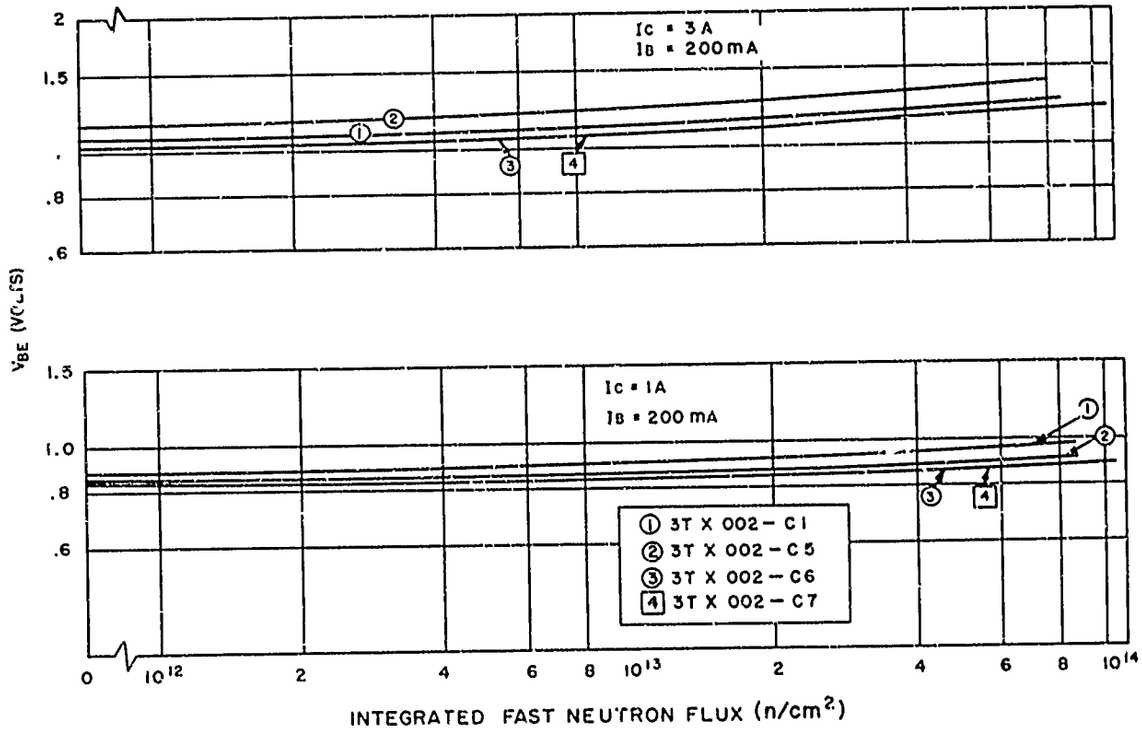


Figure II-11. Change in Emitter-Base Voltage of Clevite 3TX002 with Fast-Neutron Irradiation

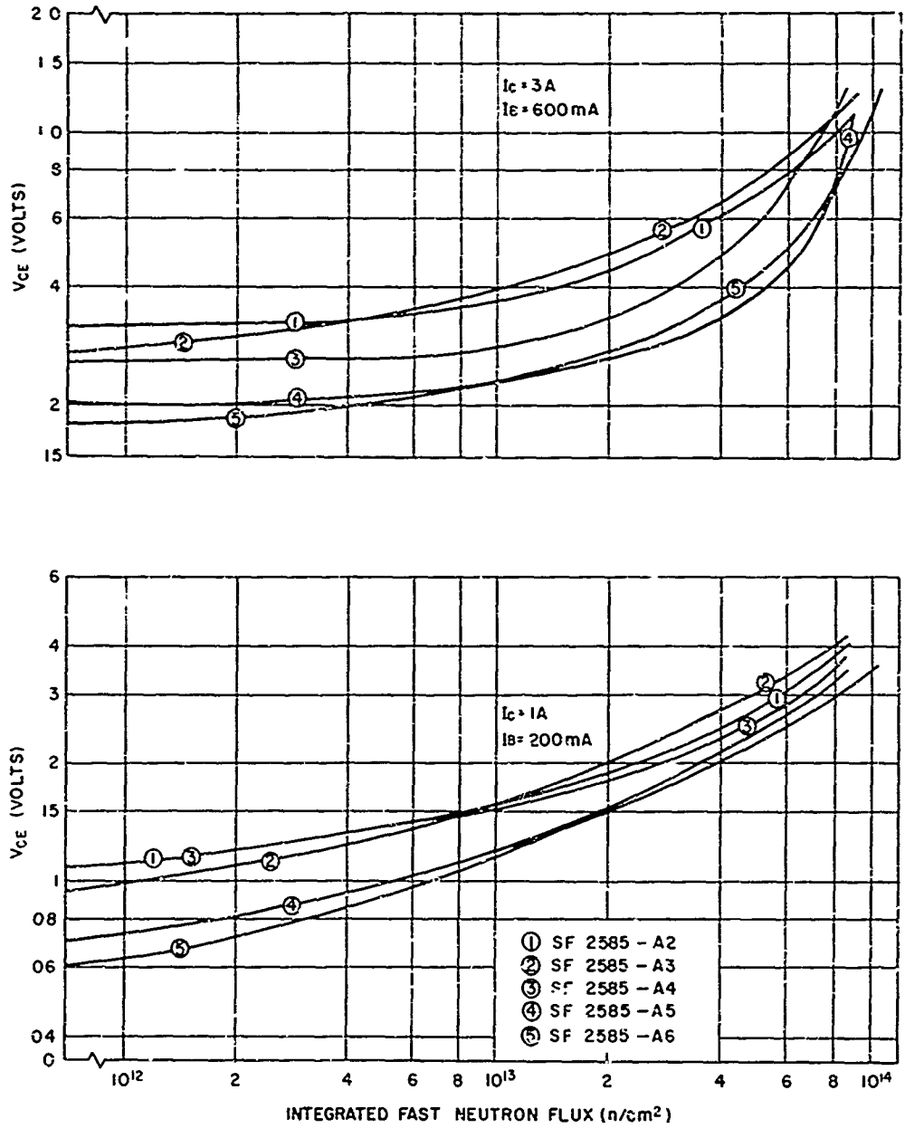


Figure II-12. Increase in Collector-Emitter Saturation Voltage of Motorola SF2585 with Fast-Neutron Irradiation

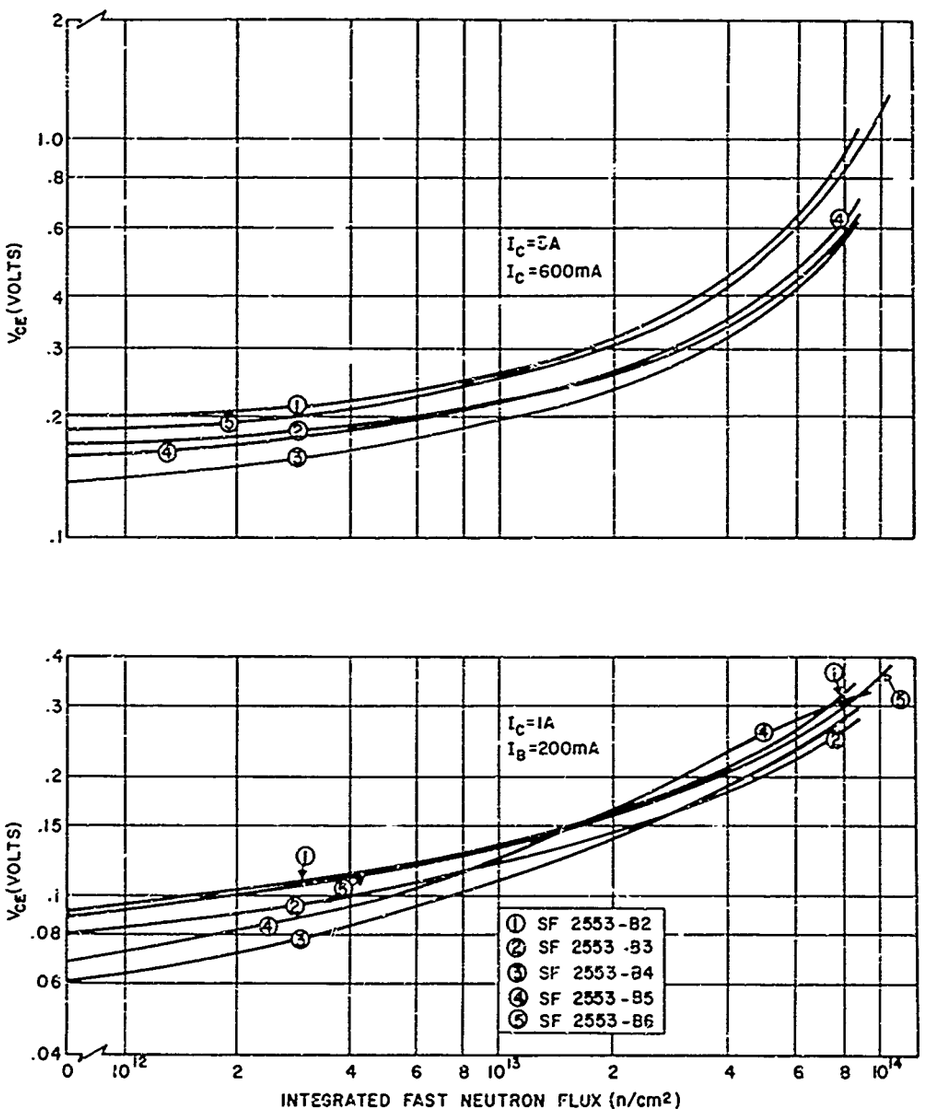


Figure II-13. Increase in Collector-Emitter Saturation Voltage of Motorola SF2553 with Fast-Neutron Irradiation

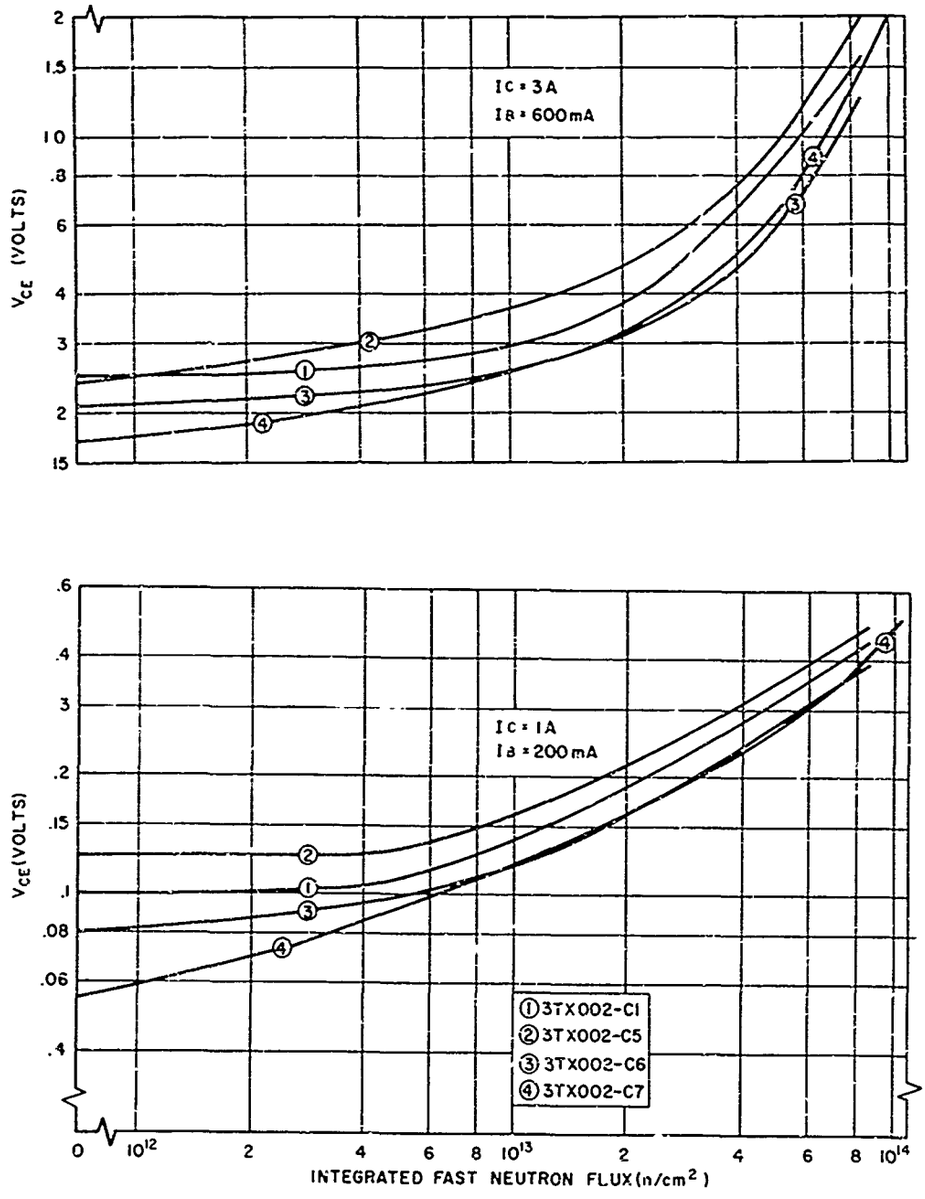


Figure II-14. Increase in Collector-Emitter Saturation Voltage of Clevite 3TX002 with Fast-Neutron Irradiation

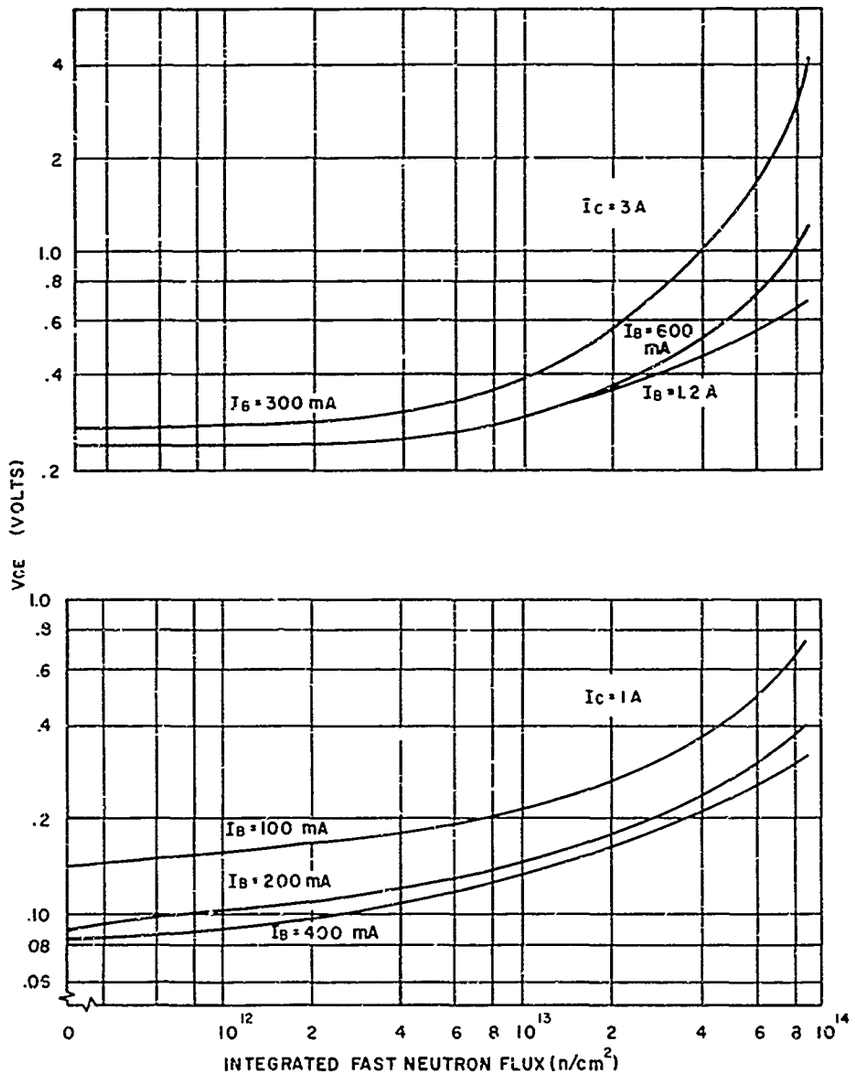


Figure II-15. Average $V_{CE(SAT)}$ versus Fast-Neutron Flux for Motorola SF2585

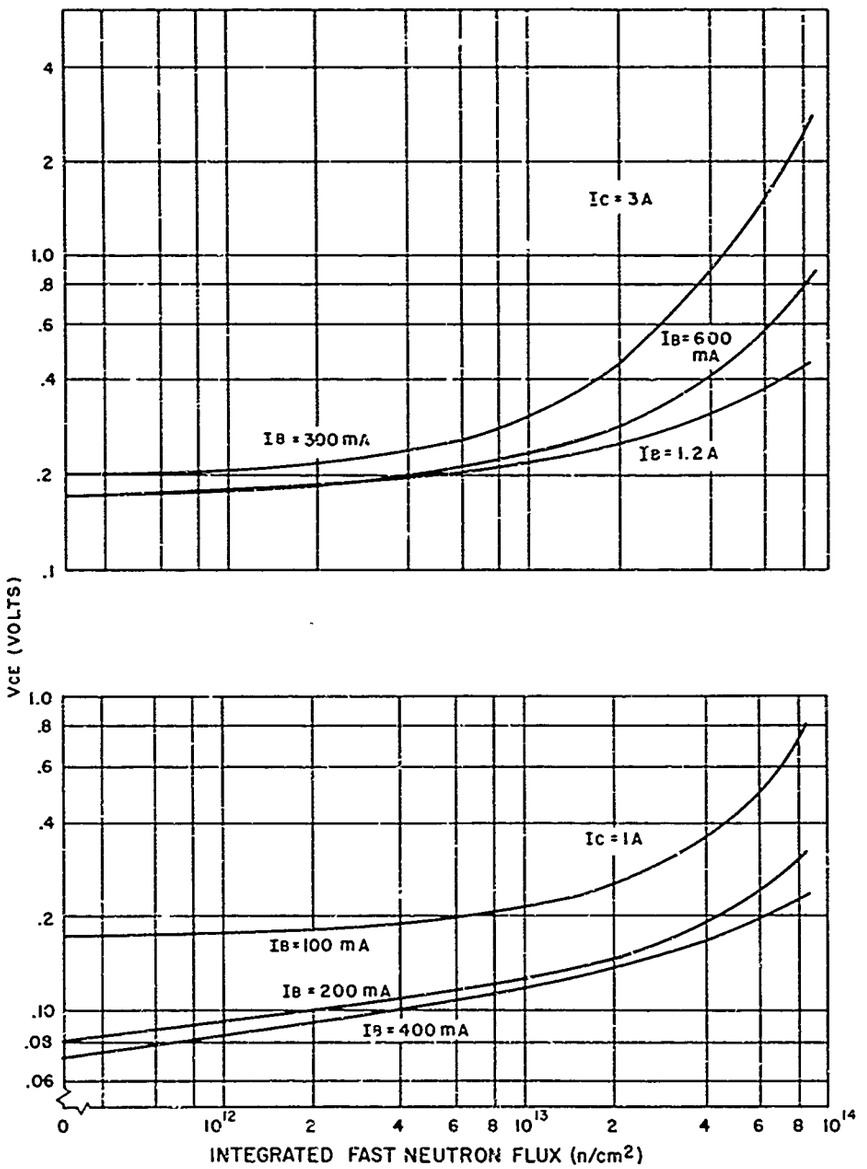


Figure II-16. Average $V_{CE(SAT)}$ versus Fast-Neutron Flux for Motorola SF2553

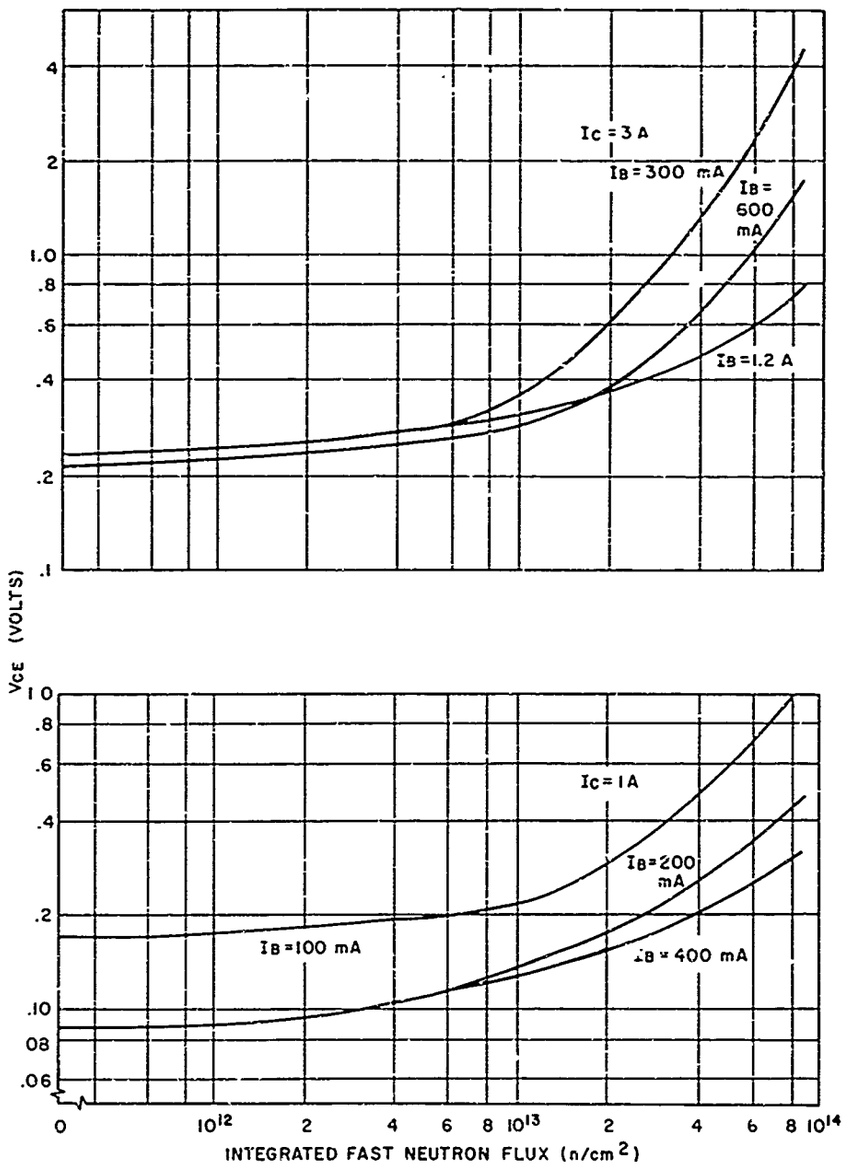


Figure II-17. Average $V_{CE(SAT)}$ versus Fast-Neutron Flux for Clevite 3TX002

The relationship of the saturated region of transistor operations to the active and "off" regions can be seen from the collector characteristics shown in Figure II-18.

When the transistor is operating in the saturated region, both junctions become forward biased. For an NPN transistor, the emitter will emit electrons into the base and the majority of these electrons will be collected by the collector, thus producing a normal transistor. Similarly the collector, since its junction is forward biased, will emit electrons into the base, some of which are collected by the emitter. This is referred to as an inverted transistor and its appropriate alpha is α_I, inverted alpha. Thus the saturated transistor becomes, in essence, a superposition of a normal transistor and an inverted transistor. The two transistors do not interact with each other to the first order because of the linearity governing the flow of carriers across the base. The voltage drop across the junctions of the saturated transistor becomes the difference of the two forward-biased junction voltages. The superposition of the two transistors is shown in Figure II-19.

A review of the literature shows that Ebers and Moll[17] have developed an equivalent circuit of the saturated transistor in which the normal and inverted transistors are at the same physical location. They have obtained an expression relating the collector and emitter junction voltages to the operating currents for a low-level current model. A derivation of this relationship, including the modifications which have to be made for high-level currents, is given in Appendix B.

Moll[18] has extended this model from one dimension to two dimensions for transistor structures in which the collector junction area is larger than the emitter junction area. He considers the effect of a lateral base current producing a potential gradient along the collector and emitter junctions. As will be shown later, this effect tends to separate the physical locations of the normal and inverted transistor.

Other authors, such as Rudenberg,[19] have considered transistors in which the collector junction area extends under the metal contact to the base region. This is the geometry that exists in the epitaxial transistor. However, these authors have assumed that the collector region of the transistor is an equipotential surface since it is assumed relatively thick and heavily doped compared with the base region, and therefore the lateral resistance of the collector is negligible.

For epitaxial transistors in which the width of the high-resistivity epitaxial layer cannot be considered small compared to the spacing between the emitter and the metal contact to the base, the lateral resistance of the collector region cannot be neglected. The problem to be considered in this section is one of presenting a model of the epitaxial transistor operating in the saturated region, including the effects of a lateral voltage gradient in the collector region. From this model, the effects of operating current and fast-neutron irradiation on the collector-emitter saturation voltage will be discussed.

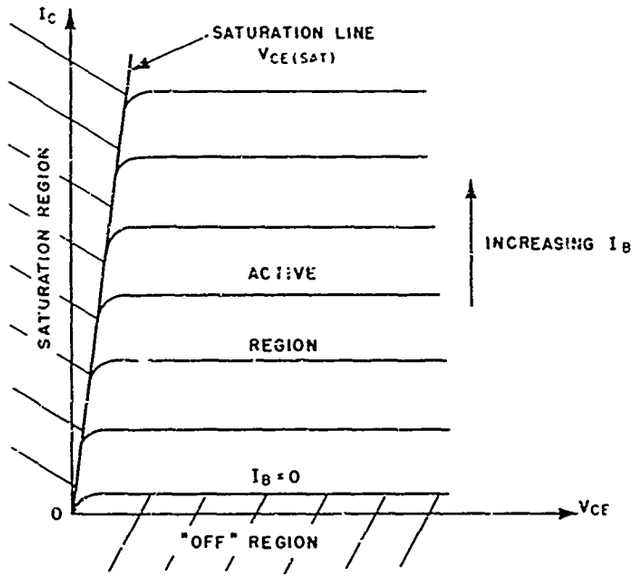


Figure II-18. Collector Characteristics for NPN Common-Emitter Transistor

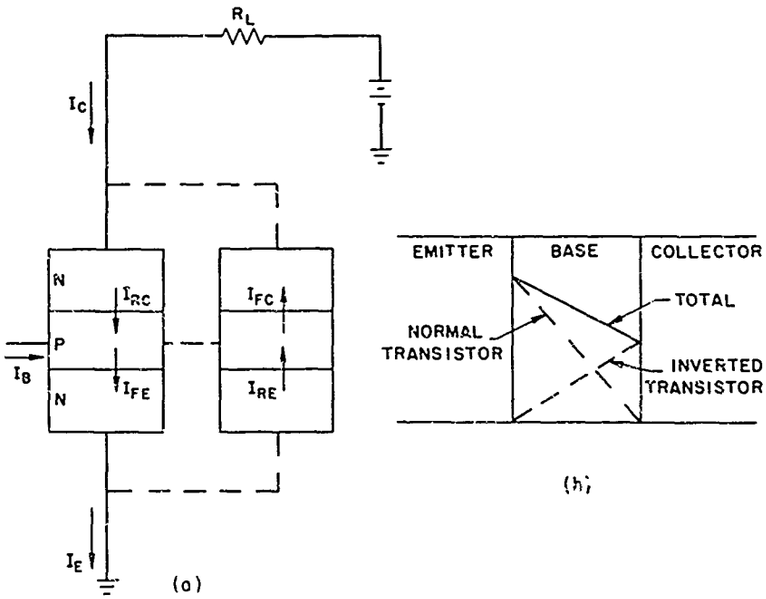


Figure II-19. Saturated Transistor Showing Superposition Which Occurs (a) and Minority-Carrier Concentration in the Base Region (b)

As mentioned in Paragraph 2, the emitter region of an epitaxial transistor is usually interleaved with the metal contacts to the base. In order to simplify analysis and still retain the important physical dimensions of the actual transistor, a physical model consisting of one emitter strip and one base contact strip will be used for analysis. The area of the emitter strip is the same as the combined emitter area of the actual transistor. The emitter-base perimeter and the spacing between the emitter region and the base contact of the model and actual transistor are the same.

The following assumptions are made in the analysis of the model and the effects of irradiation:

1. The junctions are assumed planar and parallel.
2. Voltage drops in the emitter region and collector substrate are negligible.
3. Surface recombination is neglected.
4. Step junctions are assumed except in particular cases in which there would be a significant difference in properties of a diffused junction and a step junction.
5. Changes in conductivity of the semiconductor regions with irradiation are negligible for the fast-neutron flux levels considered in this investigation.

The validity of these assumptions is discussed in Appendix A.

The physical model and equivalent circuit are shown in Figures II-20 and II-21. V_{JC} and V_{JE} are the most forward-biased collector and emitter voltages, respectively. The base and collector currents, I_B and I_C, are the independent variables for a saturated transistor in a common-emitter circuit configuration. The current components within the transistor can be written in terms of I_C and I_B as

$$I_E = I_{FE} - \alpha_I I_{FC} = I_C + I_B \tag{10}$$

$$I_C = \alpha_N I_{FE} - I_{FC} \tag{11}$$

Solving the above equations yields

$$I_{FE} = \frac{(1 - \alpha_I) I_C + I_B}{1 - \alpha_N \alpha_I} \tag{12}$$

$$I_{FC} = \frac{\alpha_N I_B - (1 - \alpha_N) I_C}{1 - \alpha_N \alpha_I} \tag{13}$$

The voltage V_{CE} can be written as

$$V_{CE} = V_{JE} - V_{JC} - V_D + V_R + I_C R'_C \tag{14}$$

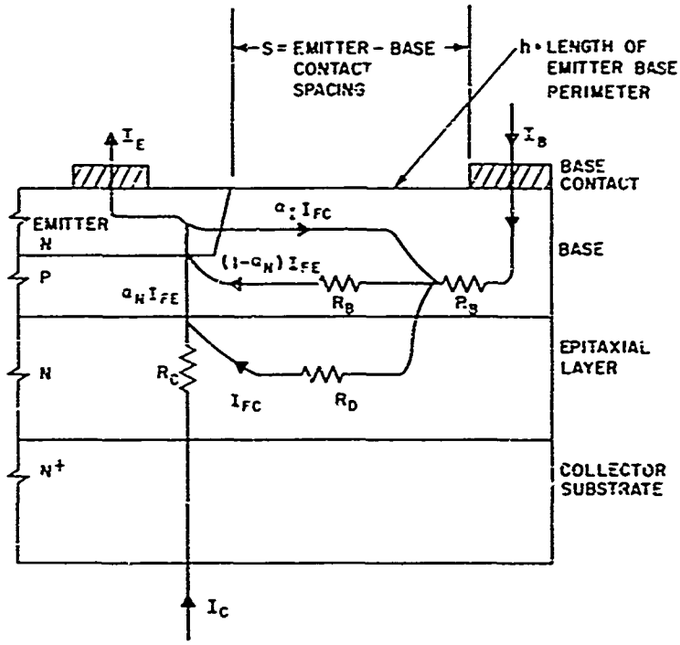


Figure II-20. Physical Model and Currents in Saturated Transistor

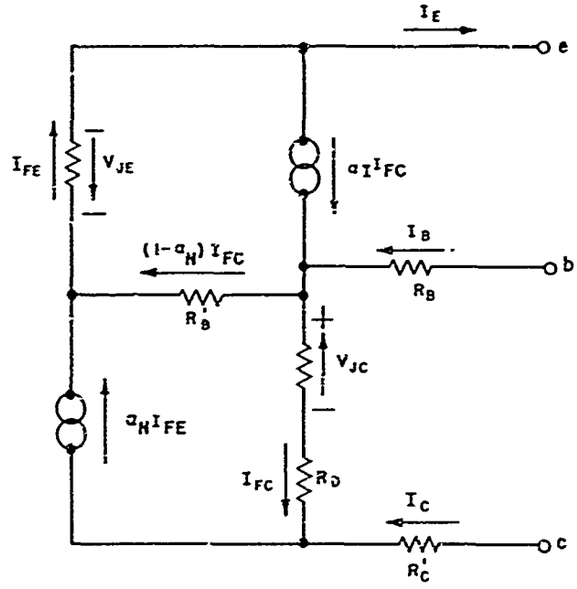


Figure II-21. Equivalent Circuit of Saturated Transistor

where V_R is the lateral voltage drop in the base region and V_D is the lateral voltage drop in the collector region. V_R and V_D can be calculated from the model as

$$V_R = R'_B \left(1 - \alpha_N\right) I_{FE} \tag{15}$$

$$V_R = \frac{R'_B \left(1 - \alpha_N\right)}{1 - \alpha_N \alpha_I} \left[\left(1 - \alpha_I\right) I_C + I_B\right] \tag{16}$$

and

$$V_D = R_D \, I_{FC} \tag{17}$$

$$V_D = \frac{R_D}{1 - \alpha_N \alpha_I} \left[\alpha_N I_B - \left(1 - \alpha_N\right) I_C\right] \tag{18}$$

Equations (14) through (18), along with Figure II-21, appear to give a rather simple model for calculating V_{CE} of the saturated transistor. This model is sufficient for most alloy and nonepitaxial diffused transistors. However, for the epitaxial transistor, the calculation of the individual terms of Equation (14) becomes a rather complicated process.

Before these terms can be evaluated, the physical location of the normal and inverted transistor must be determined. First consider the normal transistor involving the currents I_{FE}, $\alpha_N I_{FE}$, and $(1 - \alpha_N) I_{FE}$. Due to the large value of sheet resistance in the narrow base region under the emitter, the base current $(1 - \alpha_N) I_{FE}$ encounters considerable resistance as it flows laterally in the base. This resistance gives rise to a potential gradient underneath the emitter-base junction. Since emission varies as exp (40 V/n), $(1 \leq n \leq 2)$, the emitter current is effectively localized on the outer periphery of the emitter. As a result of this edge-crowding effect, emitter and collector current densities can be quite high in a small area under the emitter periphery. Hauser[20] has shown that the effective area of the emitter can be expressed as

$$A' = \frac{2\,kT}{q} \frac{h^2 \, W_{b1}}{I_{FE}^2 \left(1 - \alpha_N\right) \rho'_{b1}} \text{ cm}^2 \tag{19}$$

where ρ'_{b1} is the resistivity of the narrow base region under the emitter, taking into account conductivity modulation as discussed in Appendix C. In order to increase the effective area, the emitter-base perimeter (h) should be as large as possible for a given emitter area, thus the reason for the interleaving of emitter area and base contacts.

The location of the inverted transistor is also dependent upon the potential gradient produced by the lateral base current. The potential gradient in the base tends to crowd the inverse emission of electrons by the collector, and the injecting of holes into the collector by the base, into a localized area underneath the base contact. Thus the normal and inverted transistor would be separated by the distance from the emitter region to the base metal contact. However, the electrons needed for inverse emission and for recombination with holes injected into the collector must come from the collector area underneath the emitter. This lateral flow of electrons in the collector will produce a potential gradient which makes that part of the collector junction nearest the emitter the most forward-biased part. Thus the lateral potential gradient in the base and the lateral potential gradient in the collector have opposite effects in determining where the inverted transistor action occurs.

The dependence of the location of the inverted transistor upon the relative magnitudes of V_R and V_D can be demonstrated by considering Figures II-22 through II-27. The assumption $\alpha_I = 0$ can be made without affecting the principle demonstrated.

When the base current I_B is just large enough to drive the transistor into saturation, $I_B \approx (1 - \alpha_N)I_E$ and $I_{FC} << I_B$, making V_D negligible compared to V_R. In this condition the collector is an isopotential surface and the potential gradient in the base crowds the inverted transistor into a localized area under the base contact. The voltage distribution in the transistor and the paths of hole and electron flow for this condition are given in Figures II-22 and II-23. In the voltage distribution diagram the collector junction bias gradient is given by the voltage between the dotted lines representing V_R and V_D. The relative magnitudes of the voltage components in Figure II-22 are exaggerated in order to demonstrate the gradient in the collector junction bias. In Figure II-23 I_{RP} is that current due to holes being injected into the collector from the base, and I_{RN} is due to electrons injected into the base from the collector. The relative magnitudes of I_{RP} and I_{RN} are determined by the injection efficiency of the collector. It should be noted that in the figures showing flow paths of carriers, the lateral flow of electrons in the collector is shown as being in the epitaxial layer, whereas, in reality, the lateral flow would be distributed between the epitaxial layer and the collector substrate. The distribution would depend upon the ratio of the width of the epitaxial region to the lateral distance that the electrons must flow.

As the base current is increased and the collector current is held constant, the lateral current in the base region will increase only a small amount equal to $(1 - \alpha_N) I_B$. The majority of the increase in base current will flow laterally in the collector. Thus V_D is increasing at a much faster rate than V_R. The rate can be expressed as

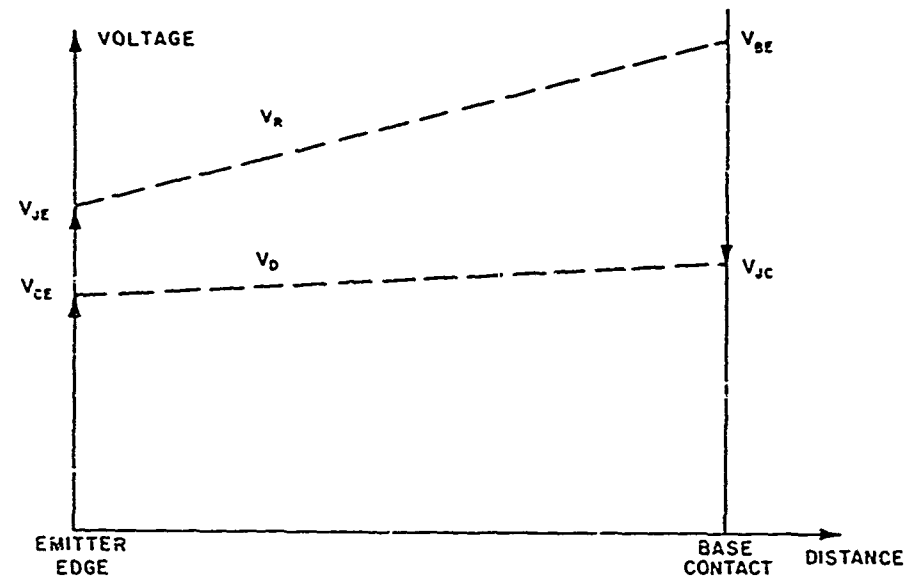


Figure II-22. Voltage Distribution in Transistor for $V_D \ll V_R$

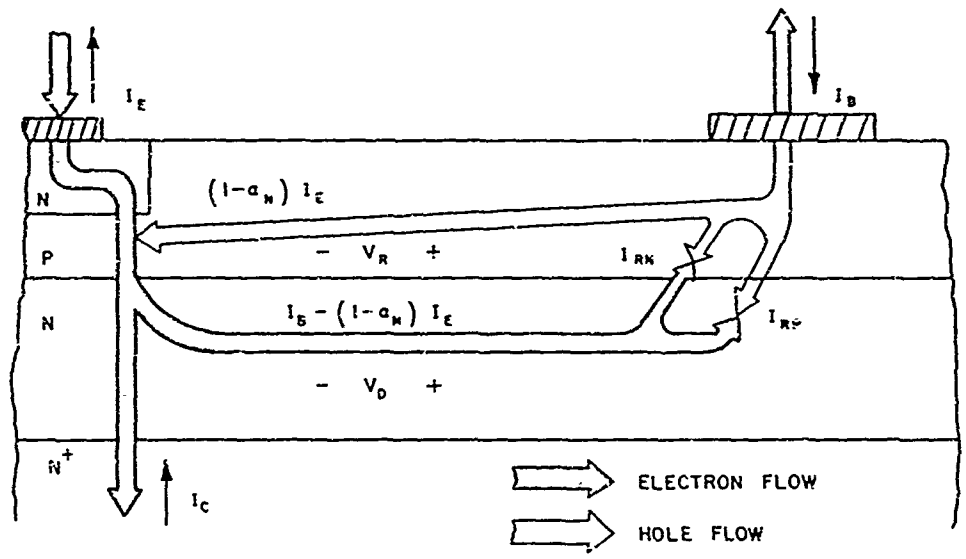


Figure II-23. Electron and Hole Flow in Transistor for $V_D \ll V_R$

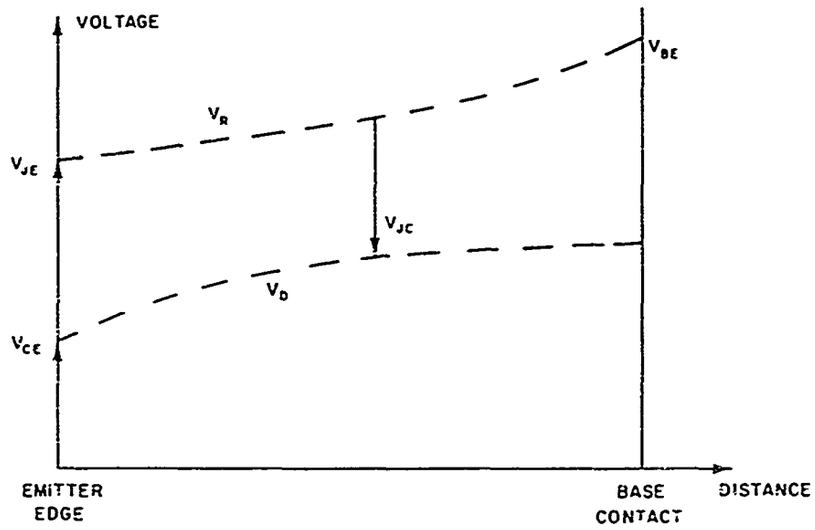


Figure II-24. Voltage Distribution in Transistor for $V_D \approx V_R$

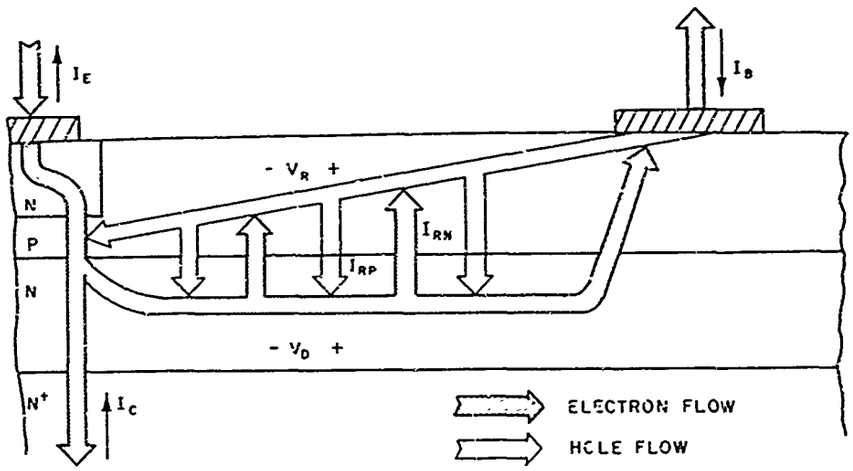


Figure II-25. Electron and Hole Flow in Transistor for $V_D \approx V_R$

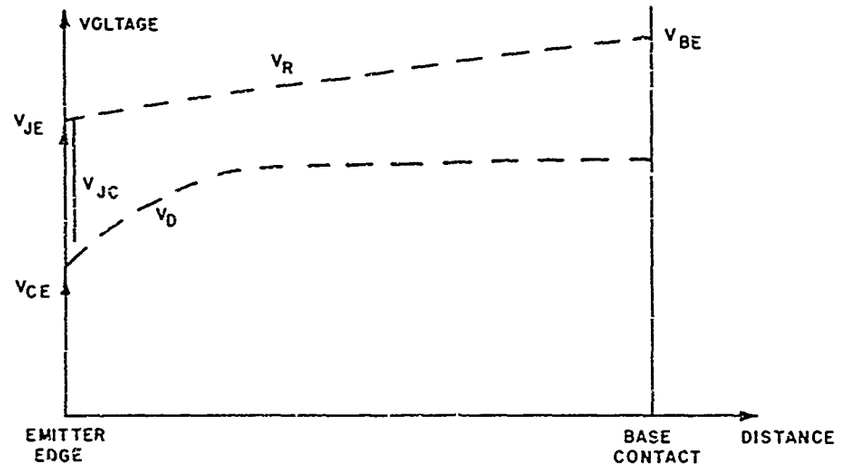


Figure II-26. Voltage Distribution in Transistor for $V_D >> V_R + I_B R_B$

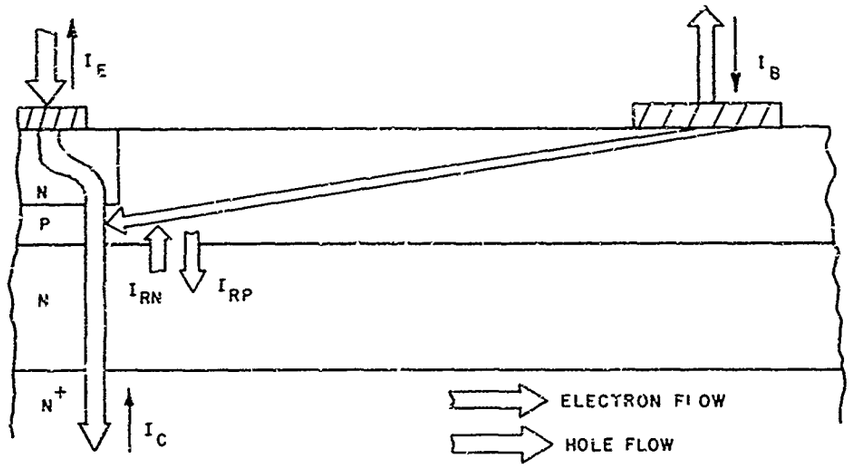


Figure II-27. Electron and Hole Flow in Transistor for $V_D >> V_R + I_B R_B$

$$\Delta V_D = \frac{\alpha_N}{1 - \alpha_N} \Delta V_R = \beta_N \Delta V_R \tag{20}$$

As the base current is increased, a value will be reached for which $V_D \approx V_R$. Under this condition, the collector junction has approximately the same forward bias at all locations, as shown in Figure II-24. The inverted transistor is now distributed over the entire junction, as shown in Figure II-25.

As the base current is increased past the value for which $V_D \approx V_R$, V_D will become larger than V_R and will tend to crowd the inverted transistor into a localized area close to the edge of the emitter, as shown in Figures II-26 and II-27. Now V_D must be larger than $V_R + I_B R_B$, since V_R has been defined by Figure II-21 to be the lateral voltage drop in the base region between the normal and inverted transistor.

Therefore it can be seen that increasing the base current beyond the minimum value required to saturate the transistor has the effect of pulling the inverted transistor from under the base contact to the edge of the emitter. Chaplin[21] has verified experimentally the dependence of the position of inverse emission by the collector upon the lateral voltage drops in the base and collector.

The lateral base resistance R'_B can be calculated from

$$R'_B = \frac{\rho_{b2}\, \ell}{W_{b2}\, h} \text{ ohms} \tag{21}$$

where ρ_{b2} and W_{b2} are the resistivity and width of that part of the base region that is not under the emitter and ℓ is the lateral distance between the normal and inverted transistor.

The calculation of R_D is not as straightforward as R'_B since the low-resistivity collector substrate affects the lateral resistance of the epitaxial layer. Champion[22] has analyzed the potential distribution in a semiconductor block with strip contacts and a conducting bottom surface. The epitaxial region can be considered as a similar model since the higher resistivity substrate behaves essentially like a metal. R_D can then be approximated by

$$R_D = \frac{\rho_{ep}\, \ell}{W_{ep}\, h}\, g \text{ ohms} \tag{22}$$

where ρ_{ep} and W_{ep} are the resistivity and width of the epitaxial layer and g is a reduction factor which depends upon the ratio ℓ/W_{ep}, as shown in Figure II-28.

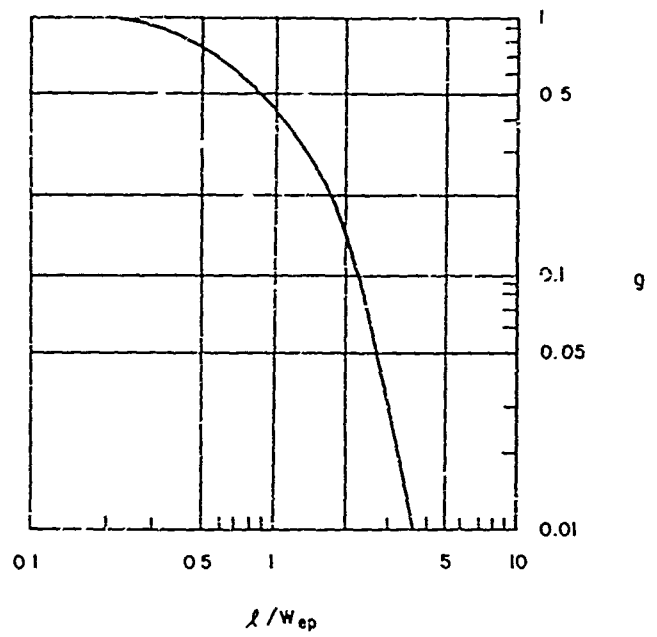


Figure II-28. Reduction Factor for Sheet Resistance of Semiconductor with Conducting Bottom Surface

It should be noted that R'_B decreases as the inverted transistor is brought closer to the normal transistor due to the decrease in ℓ, but that R_D will not decrease proportional to ℓ but may actually increase due to the increase of g.

Considering the model presented in Figure II-21, the approximate base current needed for $V_D \approx V_R$ can be found by solving for the value of I_B which satisfies the following equation for a given I_C:

$$\frac{(1 - \alpha_N)\, I_{FE}\, \rho_{b2}}{W_{b2}} = \frac{I_{FC}\, \rho_{ep}}{W_{ep}}\, g \tag{23}$$

Using the expressions for I_{FE} and I_{FC} given by Equations (12) and (13), the above equation can be written in terms of I_C and I_B as

$$I_B \left[\frac{\alpha_N\, \rho_{ep}}{W_{ep}}\, g - \frac{(1 - \alpha_N)\, \rho_{b2}}{W_{b2}} \right] = I_C \left[\frac{(1 - \alpha_I)(1 - \alpha_N)\, \rho_{b2}}{W_{b2}} + \frac{(1 - \alpha_N)\, \rho_{ep}}{W_{ep}} \right] \tag{24}$$

The necessary value of I_B needed before the inverted transistor can be considered to be crowded into a localized area near the emitter edge will, of course, be larger than the value calculated from the above equation, but the equation can be used to determine a minimum I_B needed to move the physical location of the inverted transistor toward the normal transistor.

Returning to Equation (14), the junction voltages V_{JC} and V_{JE} will now be considered. As shown in Appendix B, the Ebers and Moll expression for $V_{JC} - V_{JE}$ is

$$V_{JC} - V_{JE} = \frac{nkT}{q} \ln \frac{\alpha_I \left\{ 1 - \left[I_C/I_B \right] \left[(1 - \alpha_N)/\alpha_N \right] \right\}}{1 + (I_C/I_B)\ (1 - \alpha_I)} \text{ volts} \qquad (25)$$

where $1 \leq n \leq 2$ for silicon. For the operating conditions examined for the particular transistors used in this investigation, n will not be much greater than one. As pointed out in Appendix B, the above expression is not completely correct for this analysis but it will be used since $V_{JC} - V_{JE}$ will be shown to be only a small contributor to V_{CE} after irradiation.

An examination of Equation (25) shows that $V_{JC} - V_{JE}$ increases as α_N or α_I decreases. The dependence of α_N upon fast-neutron irradiation can be seen from an approximate equation for α_N of a NPN transistor,

$$\alpha_N \approx \left(1 - \frac{W_{b1}^2}{2\, D_{nb}\, \tau_{nb}} \right) \left(1 + \frac{\rho_e}{\rho_{b1}} \frac{W_{b1}}{L_{pe}} \right)^{-1} \qquad (26)$$

The first term on the right-hand side of the above equation is known as the base transport factor and will decrease as τ_{nb} decreases. It was shown in Paragraph 1 that minority-carrier lifetime decreases with fast-neutron irradiation, thus reducing the base transport factor. Combining Equations (26) and (8), the base transport factor (β_N^*) can be written in terms of the integrated fast-neutron flux (ϕ) as

$$\beta_N^* = \left[1 - \frac{W_{b1}^2}{2\, D_{nb}} \left(\frac{1}{\tau_{nbi}} + \frac{\phi}{K} \right) \right] \qquad (27)$$

where τ_{nbi} is the minority-carrier lifetime in the base before irradiation. The second term on the right-hand side of Equation (26) is known as the emitter efficiency term and it is also dependent upon minority-carrier lifetime since $L_{pe} = (D_{pe}\, \tau_{pe})^{1/2}$. However, for the low-resistivity emitter regions considered in this investigation, the emitter efficiency term will be approximately unity and will change very little with irradiation. Thus the primary change in α_N with irradiation can be attributed to the reduction in base transport factor.

A simple approximation cannot be made for α_I as was made for α_N. An exact expression must be used for the base transport factor of the inverted transistor. This is given by

$$\beta_I^* = \operatorname{sech} \frac{W_b'}{\left(D_{nb}\,\tau_{nb}\right)^{1/2}} \tag{28}$$

where W_b' is not a constant but is dependent upon the physical location of the inverted transistor. As the inverted transistor is moved from under the base contact toward the emitter, W_b' will decrease and β_I^* will increase. The minority-carrier lifetime τ_{nb} will decrease with fast-neutron irradiation as discussed for α_N.

An expression for the injection efficiency of the collector acting as an emitter in the inverted transistor is difficult to obtain, due to the effects of the N^+ substrate, in addition to the epitaxial layer in determining the injection efficiency. If only the epitaxial layer were considered, the injection efficiency would be very low, since the concentration of holes in the base is approximately two orders of magnitude higher than the electron concentration in the epitaxial region. Therefore the current flow across the forward-biased collector junction will consist almost entirely of holes being injected into the collector from the base. However, it requires only a small density of hole current ($\approx 10 A/cm^2$) in order for the injected hole concentration to become larger than the electron concentration in the epitaxial region. In order to maintain charge neutrality, electrons must be injected into the epitaxial region from the N^+ substrate. This increases the electron concentration in the epitaxial region, thus increasing the injection efficiency of the collector. The injection efficiency is no longer a function of the impurity doping but rather is a function of injected hole current. An approximate expression for collector injection efficiency (γ_I) as a function of injected hole current has been derived in Appendix D as

$$\gamma_I \approx \left[1 + \frac{q\,D_{pc}^2\,W_{b2}\,P_{po}}{D_{nb}\,L_{pc}^2\,\tanh^2\left(W_{ep}/L_{pc}\right)\,J_p}\right]^{-1} \tag{29}$$

Since J_p is also a function of γ_I, an exact value for γ_I is difficult to calculate, but Equation (29) does show the dependence of γ_I upon J_p, the minority carrier diffusion length (L_{pc}), and the widths of the transistor regions. Typical values calculated from the above equation and from measurements of a_I of the transistors used in this investigation give $a_I \sim 0.6 - 0.9$ before irradiation for reasonable values of J_p.

The last component of V_{CE} which must be considered is $I_C R_C'$. The resistance R_C' must include two important effects: current crowding and conductivity modulation. Due to the crowding of the current toward the emitter edge, a relatively large density of electrons enters the collector in a narrow region beneath the emitter edge. These electrons will spread out in the collector body, but in a region close to the collector junction there may be a considerable $I_C R_C'$ drop due to the high resistivity of the epitaxial layer. On the other hand, this IR drop can be

reduced by an order of magnitude or more if a hole current is injected into the epitaxial region at a distance of less than a hole-diffusion length from the place where the electrons enter the collector. Conductivity modulation is discussed in Appendix C, where it is shown that the effect of the hole current is to reduce the resistivity of the epitaxial region from ρ_c to ρ'_c where

$$\rho'_c = \frac{\rho_c}{1 + \left(J_p W_{ep}^2 / 2q D_{pc} N_D\right)} \tag{30}$$

where N_D is the impurity concentration in the epitaxial layer and W_{ep} is the width of the epitaxial layer. If the diffusion length (L_{pc}) of holes in the epitaxial layer becomes less than W_{ep}, then L_{pc} should be used in place of W_{ep} in Equation (30), and only a distance of L_{pc} from the junction can be considered to be conductivity-modulated. However, even when $L_{pc} << W_{ep}$, a considerable reduction in the IR drop can be achieved by increasing J_p, since the majority of the IR drop would be expected to occur in a region close to the collector junction. The spreading of the electrons away from the narrow region under the emitter edge will also produce a lateral voltage drop which helps to hold the inverted transistor close to the emitter edge. The motion of the carriers is shown in Figure II-29.

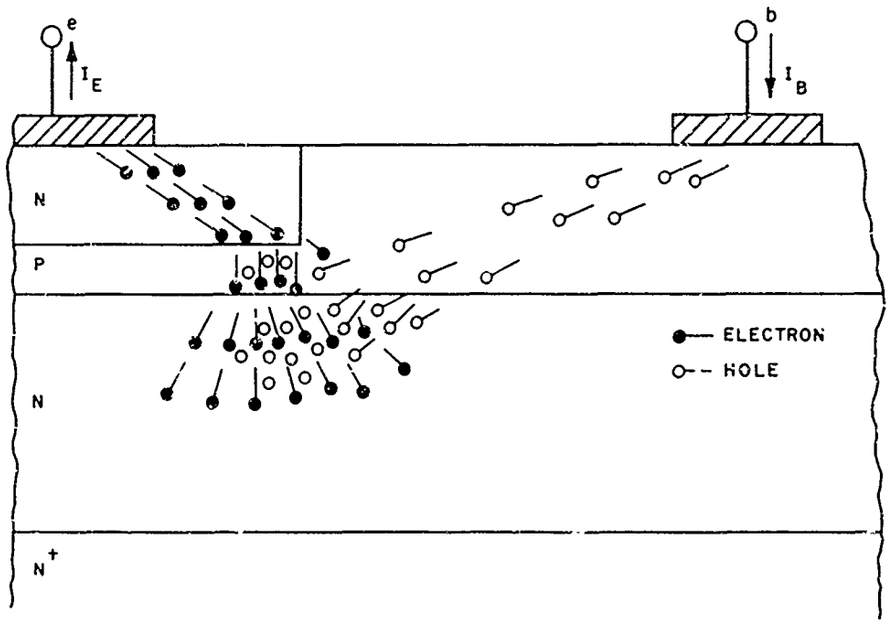


Figure II-29. Motion of Holes and Electrons in Saturated Transistor

The resistance of the epitaxial region can be determined by

$$R'_C = \frac{\rho'_{ep} W_{ep}}{A'} = \frac{\rho_{ep} W_{ep}}{A' \left[1 + \left(J_p W_{ep} / 2q D_{pc} N_D\right)\right]} \text{ ohms} \qquad (31)$$

where J_p is the injected hole density and A' is the cross-sectional area of the current path through the epitaxial region. The value of A' is difficult to determine due to the spreading of the carriers as they travel from a small area at the collector-base junction toward the collector substrate. A maximum value of R'_C can be calculated by using the area defined by Equation (19). It would be expected that the R'_C calculated using this area would be several times as large as the actual value of R'_C.

The analysis of the epitaxial transistor operating in the saturated condition can be summarized by dividing the saturation operation into three regions.

In Region I, the base current is just sufficient to drive the transistor into saturation. In this region, the inverted transistor is located under the base contact; thus α_I is low and R'_C is not conductivity-modulated.

Region II would be a transition region defined by the base current calculated from Equation (24). In this region the inverted transistor is distributed along the collector junction and α_I has been increased from Region I. If the lifetime for holes in the epitaxial region is sufficient, there will be some reduction in R'_C due to conductivity modulation. V_R is also reduced in going from Region I to Region II.

In Region III, the base current is large enough to crowd the inverted transistor toward the emitter, thus giving α_I its maximum value and reducing R'_C by a considerable amount since the base is injecting holes into the collector in the same area where electrons from the emitter are being injected.

Thus the greatest change in $V_{CE(SAT)}$ as a function of I_B / I_C would be seen in going from Region I to Region III. Once the transistor is operating in Region III, increasing I_B will not decrease $V_{CE(SAT)}$ significantly once the $I_C R'_C$ drop in the epitaxial layer has been reduced by conductivity modulation to a value that is small compared to the other components of $V_{CE(SAT)}$.

The effects of fast-neutron irradiation on the saturated transistor can be seen in terms of the three regions discussed above. First it will require a larger base current to saturate the transistor, since α_N is reduced with irradiation. Once the transistor has been saturated, it will require a larger base current to drive the transistor from Region I to Region III. This is due to the fact that the lateral voltage in the base will be greater after irradiation since $(1 - \alpha_N) I_{FE}$ is larger. Thus, a larger lateral voltage drop is needed in the collector in order to overcome the increased voltage drop in the base and pull the inverted transistor toward the emitter. This increase in $(1 - \alpha_N) I_{FE}$ with irradiation also causes the emitter current to be

crowded more than before irradiation, thus increasing the $I_C R'_C$ drop in the small region where the electron current enters the collector. Once the transistor has been driven into Region III after irradiation, the $I_C R'_C$ drop in the epitaxial region will be greater than before irradiation due to the increased crowding of the injected electrons and due to the reduction in the diffusion length of the holes injected into the epitaxial region.

Therefore, before irradiation, once the transistor is saturated a small increase in I_B will drive $V_{CE(SAT)}$ to a low constant value. After irradiation, a large change in $V_{CE(SAT)}$ will be observed as the base current is increased beyond the value required to drive the transistor into saturation. Also, $V_{CE(SAT)}$ will be dependent upon I_B over a much wider range of I_B than before irradiation. The minimum value to which $V_{CE(SAT)}$ can be driven after irradiation will be greater than the minimum value before irradiation. This is because the holes injected into the epitaxial region have a small diffusion length after irradiation and therefore conductivity-modulate only part of the width of the epitaxial region, whereas before irradiation the entire epitaxial region could be conductivity-modulated.

5. COMPARISON OF THEORETICAL PREDICTIONS WITH EXPERIMENTAL OBSERVATIONS

As discussed in Paragraph 1, the primary effect of fast-neutron irradiation with flux levels up to 10^{14} neutrons-cm^{-2} is to reduce the minority-carrier lifetime in the semiconductor regions of the transistor. Since the emitter efficiency of the normal transistor is assumed to remain constant with irradiation, the change in base transport factor with irradiation can be used to determine the reduction in minority-carrier lifetime in the base with irradiation. The expression for base transport factor is repeated here as

$$\beta_N^* = \operatorname{sech} \frac{W_{b1}}{\left(D_{nb}\, \tau_{nb}\right)^{1/2}} \tag{32}$$

Using the β_N's given by the data in Figures II-5 through II-7, the appropriate base widths, and $D_{nb} \approx 10$ cm^2-sec^{-1}, the above equation can be used to calculate τ_{nb} versus irradiation by assuming $\beta_N = \beta_N^*$. The results of these calculations are shown in Figure II-30. For fast-neutron flux levels greater than 10^{13} neutrons-cm^{-2}, $1/\tau_{nb}$ increases linearly with flux level as predicted by Equation (8). The values of $1/\tau_{nb}$ given in Figure II-30 will be slightly higher than the actual values, since unity emitter efficiency was assumed in calculating the values in Figure II-30. However, Figure II-30 does give an accurate method of determining values of K in Equation

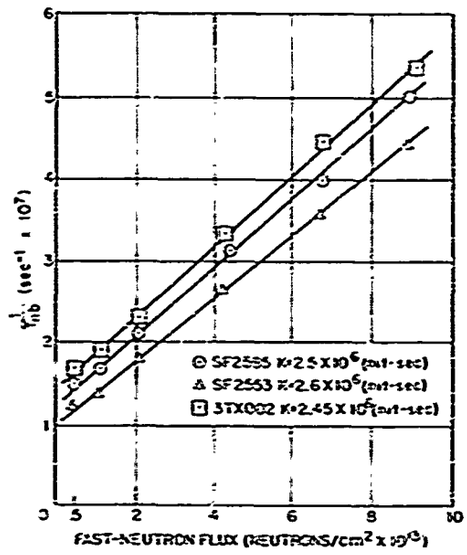


Figure II-30. Reciprocal of Lifetime of Electrons in Base versus Fast-Neutron Flux

(3). These values of K are given in the figure and are within the accepted range of values for K given in Paragraph 1.

For flux levels greater than 10^{13} neutrons-cm^{-2}, electron lifetime in the base is primarily a function of the number of radiation-induced recombination centers rather than impurity concentration. Thus the lifetime of holes in the collector can be assumed to be approximately equal to the lifetime of the electrons in the base for flux levels above 10^{13} neutrons-cm^{-2}. Values of hole lifetime in the collector have been determined from measurements of collector recovery times for these transistors before and after irradiation. These values and other pertinent data and calculations are given in Table II-2. The values given in Table II-2 vary only slightly among the different devices.

Table II-2

VALUES OF MINORITY-CARRIER LIFETIMES AND DIFFUSION LENGTHS

Parameter	Pre-Irradiation	Post-Irradiation
τ_{nb}	≈1 μsec	0.02-0.05 μsec
L_{nb}	≈30 μ	5-10 μ
τ_{pc}	2-5 μsec	0.02-0.05 μsec
L_{pc}	50-70 μ	5-7 μ

It should be noted that the measured pre-irradiation value of τ_{nb} is approximately 1 microsecond, whereas Figure II-30 indicates that the pre-irradiation value of τ_{nb} is approximately 0.1 microsecond. There are two possible reasons for this discrepancy. First, the value of $\tau_{nb} \approx 1$ microsecond given in Table II-2 is only an approximate value, since τ_{nb} is difficult to measure accurately in power transistors. Second, for irradiation levels less than 10^{13} neutrons-cm^{-2}, the assumption of unity emitter efficiency is a poor assumption since the emitter efficiency has as much effect as the base transport factor in determining β_N. As the transistors are irradiated above 10^{13} neutrons-cm^{-2}, the base transport factor becomes the dominant factor in determining β_N and the calculated and measured values of τ_{nb} are in reasonable agreement.

The values for $V_{CE(SAT)}$ before irradiation will now be considered. Equations (15) and (17) can be used to calculate V_R and V_D, using values of α_N from Figure II-8 and a measured $\alpha_I \approx 0.8$. These calculations show V_R and V_D to be on the order of 0.005-0.012 volts, which is less than the $kT/q = 0.026$ volts needed before crowding is significant. Therefore it would be reasonable to assume that the injection of holes into the collector and the injection of electrons into the base by the collector is distributed along the collector junction and will be crowded very little. However, even when the injection of holes into the collector is distributed along the collector junction, a considerable reduction in R'_C can be made, since the injected holes have a diffusion length that is long compared to the width of the epitaxial region and the spacing between the emitter and base contact. The data presented in Figures II-15 through II-17 show that, before irradiation, the largest change in $V_{CE(SAT)}$ is seen when the drive ratio (I_B/I_C) is increased from 0.1 to 0.2 with very little change in $V_{CE(SAT)}$ by increasing I_B beyond 0.2 I_C.

Since V_E and V_D are small before irradiation, $V_{CE(SAT)}$ will be determined by the junction voltage plus a small $I_C R'_C$ drop. The junction voltages can be calculated using Equation (25) and the values of α_N from Figure II-8 and $\alpha_I \approx 0.8$. An apparent R'_C can be determined from the difference in the calculated values of $V_{JE} - V_{JC}$ and the measured $V_{CE(SAT)}$. These values are given in Table II-3.

As can be seen from Table II-3, the apparent resistance R'_C is reduced by a factor of ≈ 2 when I_B is increased from 100 ma to 200 ma, and then remains essentially constant with I_B. The decrease in R'_C is due to conductivity modulation. The fact that R'_C remains constant with I_B for $I_B > 200$ ma can be explained by two factors. First, as I_B is increased, the inverse injection efficiency (γ_I) of the collector increases. Thus an increase in I_B once $I_B > 200$ ma causes a large increase in the number of electrons injected into the base by the collector, but only a small increase in the number of holes injected into the collector. Second, a resistance of 0.04 - 0.07 ohms may be due to a resistance other than R'_C, such as lead or contact resistance.

Table II-3

PRE-IRRADIATION $V_{CE(SAT)}$ AND CALCULATED COMPONENTS

Device	I_C (amp)	I_B (amp)	Calculated $V_{JE} - V_{JC}$ (volts)	Measured $V_{CE(SAT)}$ (volts)	Apparent R'_C (ohms)
SF2585	1	0.1	0.05	0.15	0.10
	1	0.2	0.037	0.09	0.053
	1	0.4	0.025	0.083	0.05[illegible]
	3	0.3	0.052	0.27	0.072
	3	0.6	0.037	0.25	0.071
	3	1.2	0.025	0.25	0.072
SF2553	1	0.1	0.052	0.17	0.118
	1	0.2	0.037	0.08	0.043
	1	0.4	0.025	0.07	0.045
	3	0.3	0.052	0.20	0.050
	3	0.6	0.037	0.17	0.044
	3	1.2	0.026	0.17	0.046
3TX092	1	0.1	0.051	0.17	0.119
	1	0.2	0.036	0.09	0.054
	1	0.4	0.024	0.09	0.060
	3	0.3	0.051	0.23	0.060
	3	0.6	0.036	0.22	0.061
	3	1.2	0.025	0.23	0.077

Now consider $V_{CE(SAT)}$ after irradiation. Figure II-8 shows the reduction in α_N. Also, α_I can be assumed to be reduced to approximately 0.3 - 0.4 by considering the reduction in lifetime for holes in the epitaxial region and the reduction in inverse base transport factor, taking into account the fact that electrons in the inverted transistor have a longer base path to travel than electrons in the normal transistor. The value of $\alpha_I \approx 0.3 - 0.4$ is when the inverted transistor is within 10 microns of the emitter. At low base currents, when the inverse emission of the electrons by the collector occurs underneath the base contact, α_I will be much lower ($\alpha_I \approx 0.03$).

When $I_B \approx I_C/\beta_N$, V_R can be calculated from Equation (15) with $I_{FE} = I_E$. This gives values of $V_R = 0.07$ for $I_E = 1.1$ amperes, and $V_R = 0.21$ for $I_E = 3.3$ amperes. Thus the inverse emission will be crowded under the base contact when the transistor starts into the saturated condition. Values of I_B/I_C required to spread the inverse emission along the entire collector junction can be calculated using Equation (24). These values of drive ratio are given in Table II-4.

Table II-4

CALCULATED VALUES OF DRIVE RATIO REQUIRED TO SPREAD OUT INVERSE EMISSION AFTER IRRADIATION

Device	I_C (amp)	I_B/I_C Required to Spread Out Inverse Emission
SF2585	1	0.13
	3	0.14
SF2553	1	0.18
	3	0.18
3TX002	1	0.14
	3	0.17

As can be seen from Table II-4, a drive ratio of 0.1 is not sufficient to move the inverted transistor from underneath the base contact. The holes injected into the collector underneath the base contact recombine with electrons before they can reach the collector area underneath the emitter periphery. Therefore, the injected holes do not help to reduce R'_C. Since the lateral base current $(1 - \alpha_N) I_{FE}$ flowing in the base under the emitter has been increased by a factor of 4 to 5 due to the decrease in α_N with irradiation, the electrons from the emitter enter the epitaxial region in a smaller area than before irradiation. As a result of these two factors, the $I_C R'_C$ drop will be larger for a drive ratio of 0.1.

Similar data can be calculated for the transistors after irradiation. The calculations are the same as for the pre-irradiation condition and the results are given in Table II-5. Again it is assumed that $V_{CE(SAT)} \approx V_{JE} - V_{JC} + I_C R'_C$, and that the difference between V_R and V_D is large enough to cause a gradient in the collector junction bias but is only a small contributor to $V_{CE(SAT)}$. To be consistent, $V_{JE} - V_{JC}$ should also be considered negligible compared to $I_C R'_C$ since $V_{JE} - V_{JC}$ is of the same order of magnitude as V_R and V_D. However, calculations of $V_{JE} - V_{JC}$ are presented for the primary purpose of showing that $V_{JE} - V_{JC}$, even with greatly reduced α_N and α_I, is not much greater after irradiation than before irradiation. Thus $V_{JE} - V_{JC}$ cannot account for the large increase in $V_{CE(SAT)}$ with irradiation.

As can be seen from the table, the apparent R'_C is reduced considerably by increasing I_B and is dependent upon I_B over the entire range of given I_B.

Several important comparisons can be made using Table II-5. These are:

1. R'_C for the Motorola SF2553 is significantly lower than for the Motorola SF2585. The most logical explanation for this is the difference in the sheet

Table II-5

POST-IRRADIATION $V_{CE(SAT)}$ AND CALCULATED COMPONENTS

Device	I_C (amp)	I_B (amp)	Calculated $V_{JE} - V_{JC}$ (volts)	Measured $V_{CE(SAT)}$ (volts)	Apparent R'_C (ohms)
SF2585	1	0.1	0.149	0.73	0.58
	1	0.2	0.110	0.38	0.27
	1	0.4	0.096	0.33	0.23
	3	0.3	0.155	3.97	1.26
	3	0.6	0.110	1.20	0.36
	3	1.2	0.096	0.68	0.20
SF2553	1	0.1	0.170	0.72	0.55
	1	0.2	0.120	0.32	0.20
	1	0.4	0.098	0.24	0.14
	3	0.3	0.185	2.74	0.85
	3	0.6	0.120	0.84	0.24
	3	1.2	0.098	0.46	0.12
3TX002	1	0.1	0.160	1.09	0.93
	1	0.2	0.120	0.47	0.35
	1	0.4	0.096	0.31	0.21
	3	0.3	0.190	4.34	1.38
	3	0.6	0.120	1.61	0.50
	3	1.2	0.097	0.74	0.21

resistance of the base region underneath the emitter, as given in Table II-1. The base sheet resistance of the SF2553 is approximately 40% less than for the SF2585, resulting in less crowding of the electrons entering the collector and thus a lower $I_C R'_C$ drop.

2. With low drive ratios, the apparent R'_C of the 3TX002 is greater than that of the SF2585. The two devices have approximately the same sheet resistance in the base region under the emitter, but the 3TX002 has a higher resistivity in the epitaxial region and a wider epitaxial region, resulting in a larger R'_C. The difference in R'_C for the two devices would be even greater if it were not for the fact that the emitter-base perimeter (h) of the 3TX002 is more than twice that of the SF2585, thus reducing the crowding of the emitter current. As the drive ratio is increased, the R'_C for the two devices approach similar values for the same drive ratio, indicating that R'_C is now determined primarily by conductivity modulation rather than by the impurity doping in the epitaxial region.

3. The diffusion length for holes injected into the epitaxial region is ≈5μ, whereas the width of the epitaxial region is 15μ for the Motorola SF2585 and SF2553 and 20μ for the Clevite 3TX002. If R'_C were distributed uniformly over the width of the epitaxial region, only 1/4 to 1/3 of R'_C could be reduced by conductivity modulation; thus R'_C could not be changed by more than 25 to 30% by increasing I_B. However, Table II-5 shows that R'_C is decreased by a factor of 5 or more by increasing I_B. This indicates that R'_C is not distributed over the width of the epitaxial region. Instead, the majority of the $I_C R'_C$ drop occurs near the collector junction where the electron path is crowded into a small area. If this is correct, conductivity-modulating only the part of the epitaxial region near the collector junction can cause a considerable reduction in R'_C.

Comparisons (1) and (2) can be seen clearly by plotting $V_{CE(SAT)}$ data for the SF2585 and SF2553 on the same graph and by plotting data for the 3TX002 and SF2585 on the same graph. These are given in Figures II-31 and II-32. It should be noted in Figure II-32 that at I_C = 3 amp, $V_{CE(SAT)}$ for the Clevite 3TX002 transistors is equal to or less than $V_{CE(SAT)}$ for the Motorola SF2585 for fast-neutron flux levels up to ≈0.6 x 10^{13} n/cm^2, and then $V_{CE(SAT)}$ for the 3TX002 starts to increase with irradiation at a faster rate than for the SF2585. The diffusion length for holes injected into the collector can be calculated by

$$L_{pc} = \left(D_{pc}\, \tau_{pc}\right)^{1/2} \text{ cm} \tag{33}$$

Hole lifetime (τ_{pc}) can be calculated using Equation (8) repeated here as

$$\frac{1}{\tau_{pc}} = \frac{1}{\tau_{pci}} + \frac{\phi}{K} \text{ sec}^{-1} \tag{34}$$

Using

$$\tau_{pci} \approx 2 \times 10^{-6} \text{ sec}$$

$$K = 2.5 \times 10^{6} \text{ sec-n/cm}^2$$

$$\phi = 0.6 \times 10^{13} \text{ n/cm}^2$$

$$\frac{1}{\tau_{pc}} = 0.29 \times 10^{7} \text{ sec}^{-1}$$

Thus

$$\tau_{pc} = 3.5 \times 10^{-7} \text{ sec}$$

and

$$L_{pc} = 18.7 \times 10^{-4} \text{ cm} = 18.7\mu$$

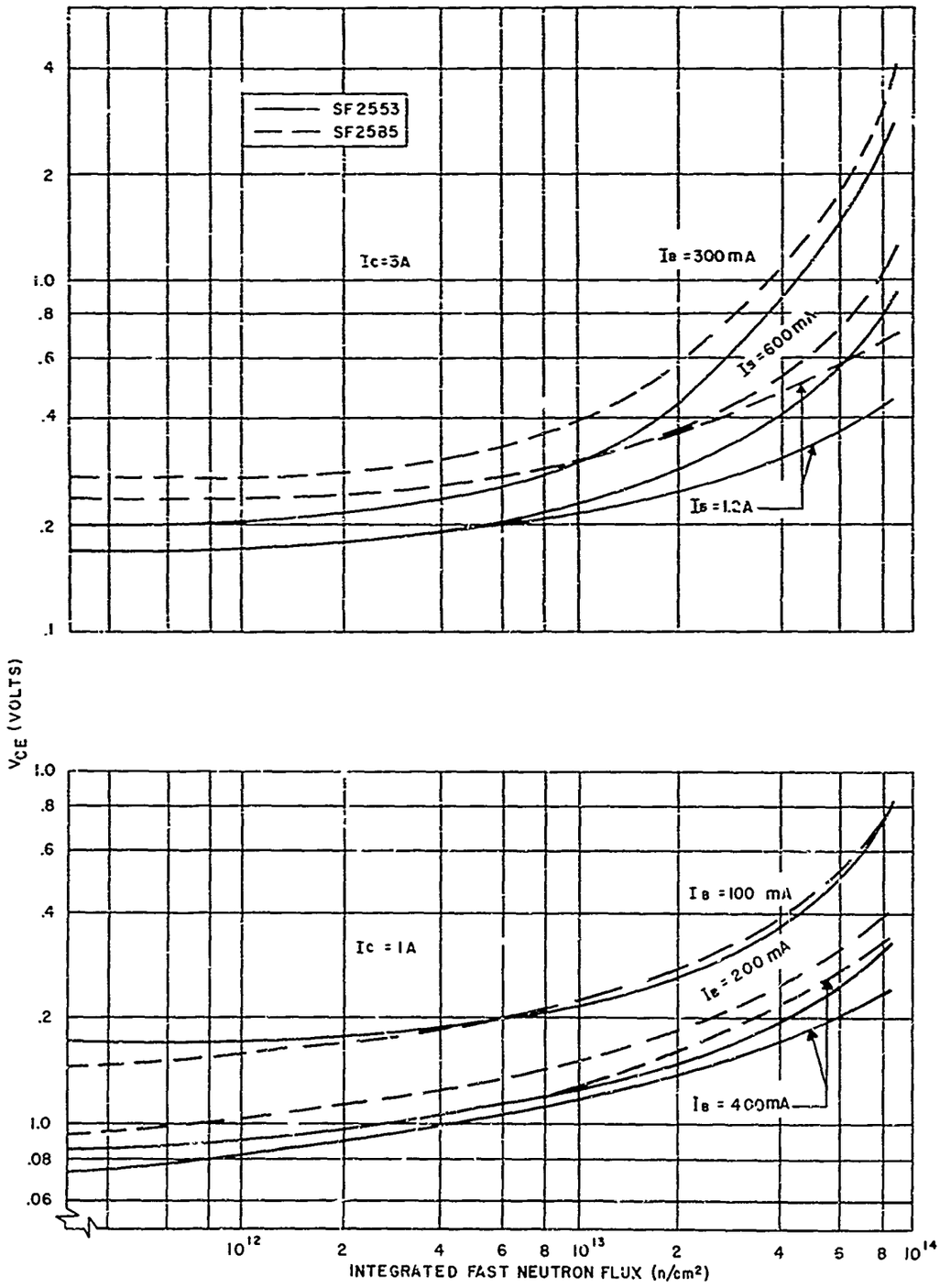


Figure II-31. Comparison of $V_{CE(SAT)}$ versus Fast-Neutron Irradiation for Motorola SF2585 and SF2553

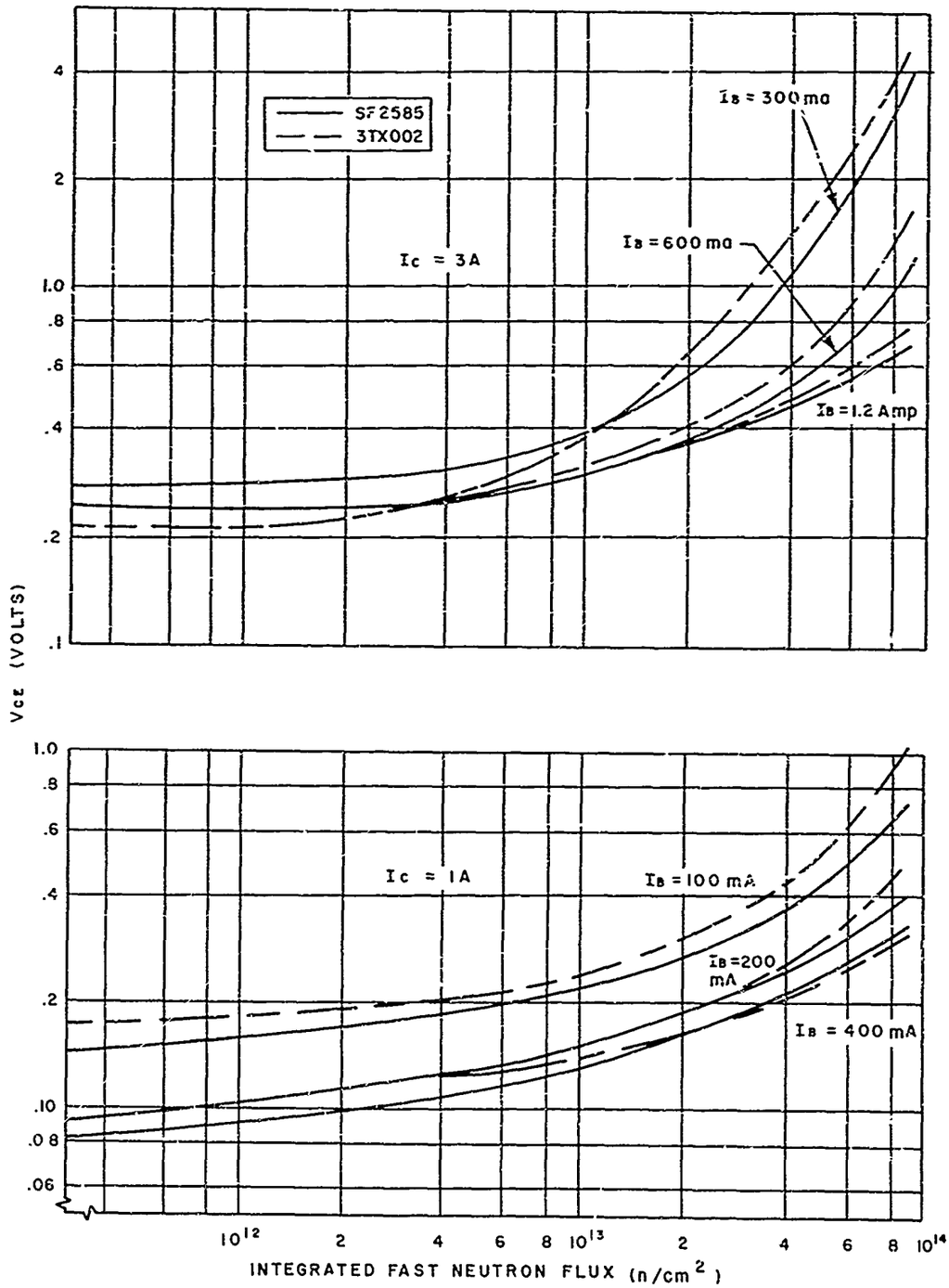


Figure II-32. Comparison of $V_{CE(SAT)}$ versus Fast-Neutron Irradiation for Motorola SF2585 and Clevite 3TX002

The width of the epitaxial region for the Clevite 3TX602 is 20[illegible]. Therefore, as the diffusion length for holes injected into the epitaxial layer becomes less than the width of the epitaxial layer, the difference in resistivity of the epitaxial layer of the two types of transistor begins to result in a significant difference in $V_{CE(SAT)}$. This is to be expected, since R'_C is then determined partly by the amount of conductivity modulation and partly by the resistivity of the epitaxial layer. This effect can be extended to all the $V_{CE(SAT)}$ measurements given in Figures II-15 to II-17 and Figures II-31 and II-32, especially for I_C = 3 amp. It can be seen that the slope of $V_{CE(SAT)}$ versus irradiation begins to increase significantly for flux levels in the range 0.6×10^{13} to 1×10^{13} neutrons-cm^{-2}. This is the same level of irradiation for which the diffusion length of holes in the epitaxial layer becomes less than the width of the epitaxial layer.

Figure II-31 shows that $V_{CE(SAT)}$ for the SF2585 is consistently higher than $V_{CE(SAT)}$ for the SF2553 even though the SF2553 has the lower α_N. Thus current gain alone is not an accurate criterion for predicting $V_{CE(SAT)}$.

Calculations of R'_C with any accuracy is difficult for two reasons. First, the effective area of the electron path through the epitaxial region varies with distance from the collector junction due to the spreading of the electrons from the small area in which they enter the collector. Second, since electrons and holes are injected into the collector over different areas, their interaction and the resulting conductivity modulation is difficult to determine with any accuracy. The relative R'_C of the devices can be approximated by calculating an R'_C assuming that the effective area is constant through the epitaxial layer and equal to the area in which the electrons from the emitter enter the collector. It is also assumed that there is no conductivity modulation, as would be expected to be the situation for a 0.1 drive ratio after irradiation. The effective area can be determined by combining Equation (19) and Appendix C Equation (5) for ρ_{b1}, resulting in

$$A'^2 - \frac{2\,kTh^2\,W_{b1}}{q\,I_{FE}^2\left(1-\alpha_N\right)\rho_{b1}}\,A' - \frac{kTh^2\,W_{b1}^2}{q\,I_{FE}\left(1-\alpha_N\right)\rho_{b1}\,D_{nb}\,N_D} = 0 \qquad (35)$$

Taking into account conductivity modulation of ρ_{b1} tends to limit the crowding of the current to a minimum area. R'_C can then be calculated from

$$R'_C = \frac{\rho_c\,W_{ep}}{A'} \qquad (36)$$

Effective areas and resistances for the devices are given in Table II-6.

Comparing Table II-6 with Table II-5, it can be seen that R'_C, assuming no spreading of the electrons in the epitaxial layer, is approximately four times that

Table II-6

CALCULATED COLLECTOR RESISTANCE AFTER IRRADIATION ASSUMING NO CURRENT SPREAD OR CONDUCTIVITY MODULATION

Device	I_C (amp)	I_B (amp)	A' (cm^2)	R'_C (ohms)
SF2585	1	0.1	1.34×10^{-3}	2.24
	3	0.3	0.58×10^{-3}	5.17
SF2553	1	0.1	1.92×10^{-3}	1.56
	3	0.3	0.64×10^{-3}	4.7
3TX002	1	0.1	4.5×10^{-3}	2.22
	3	0.3	1.68×10^{-3}	5.95

of the measured apparent R'_C. This would appear to be a reasonable reduction factor, since R'_C versus I_B indicates that the majority of R'_C is within a third to a fourth of the width of the epitaxial layer from the collector junction. Table II-6 does give the same relative values of R'_C for the different devices. It can be seen that the 3TX002 has the largest effective area due to its larger value of emitter-base perimeter, but this advantage is lost in calculating R'_C due to the higher resistivity of the epitaxial region.

One final point should be mentioned in this chapter: the shape of the collector characteristics of the transistors before and after irradiation. Before irradiation, the saturation region is defined by a definite saturation line, as demonstrated in Figure II-18. However, after irradiation, the curves become much more rounded in the vicinity of the saturation region due to the larger range of saturation voltages for different values of collector and base currents. This is illustrated by Figure II-33, which gives typical collector characteristics of the Motorola SF2585 before and after irradiation. Note that before irradiation increasing I_B once the transistor is saturated does not decrease V_{CE} for a given collector current, whereas after irradiation increasing I_B changes V_{CE} significantly for a given collector current.

6. CONCLUSIONS

The primary result of this investigation is not a rigorous formula for calculating collector-emitter saturation voltage versus fast-neutron irradiation. Instead, this investigation has shown, by means of a theoretical analysis and comparison of experimental data for different transistors, the important factors in determining the change in $V_{CE(SAT)}$ with irradiation, thereby providing a criterion for predicting

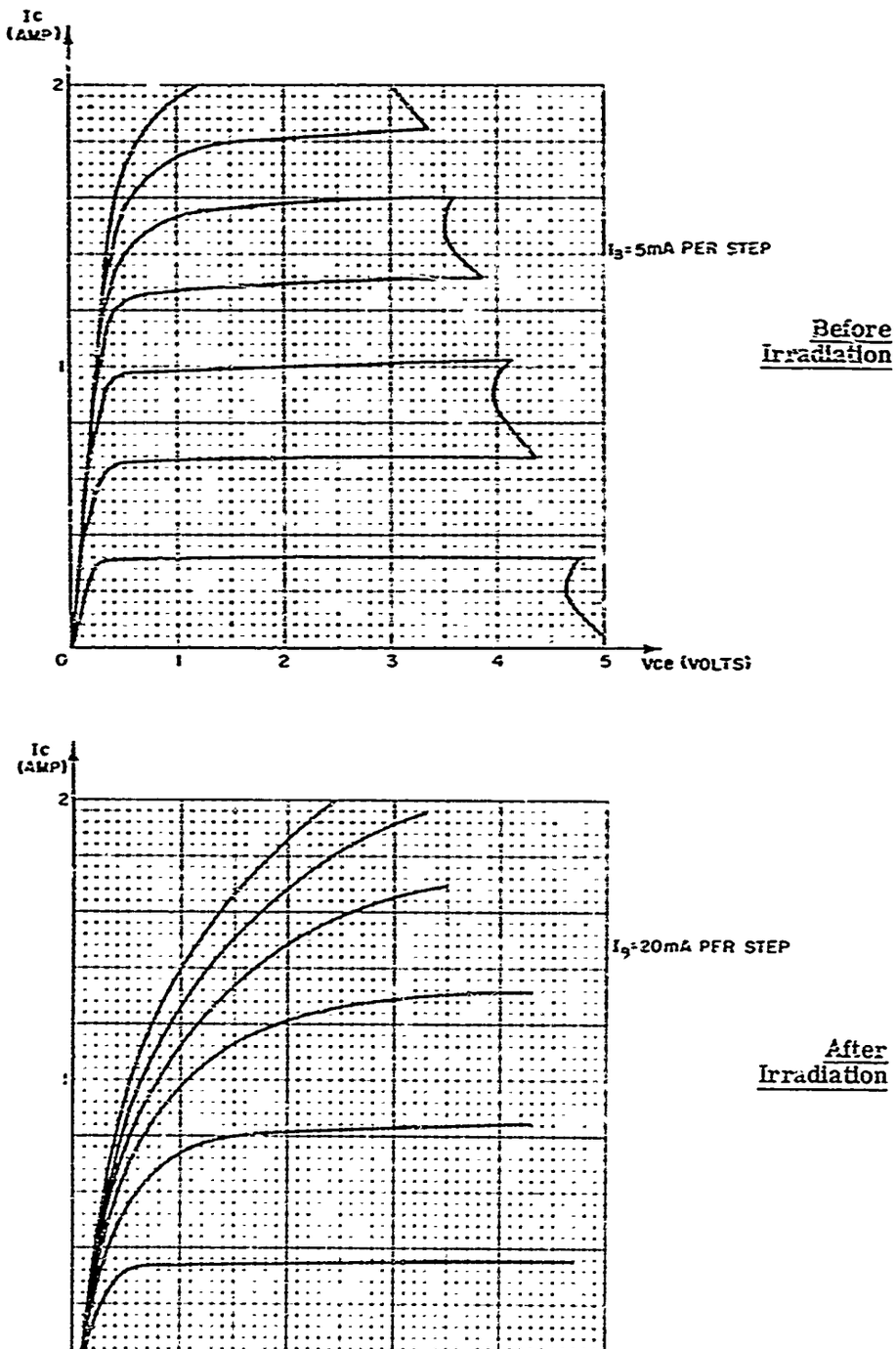


Figure II-33. Common-Emitter Collector Characteristics of Motorola SF2585 Before and After Irradiation

the relative change in $V_{CE(SAT)}$ with fast-neutron irradiation for different transistors. The results of the investigation can also be used to influence the design of switching transistors for use in radiation environments.

In Paragraph 7, the advantages of the epitaxial transistor over other types of transistors for switching applications were discussed. It was pointed out that a compromise between breakdown voltage, saturation voltage, junction capacitance, and switching time was eliminated by the use of an epitaxial layer in the collector. This is true before irradiation since the experimental results showed that the collector resistance was small and the $I_C R'_C$ drop did not contribute significantly to $V_{CE(SAT)}$. However, it was shown in the analysis and data that after fast-neutron irradiation with integrated flux levels greater than 10^{13} neutrons-cm^{-2}, it can no longer be assumed that the epitaxial transistor has negligible collector resistance. On the contrary, it was shown that the $I_C R'_C$ drop was the dominant factor in determining $V_{CE(SAT)}$ after irradiation. From this investigation, several methods of reducing the effective collector resistance can be seen. These are listed below:

1. The most obvious method is to increase the impurity doping in the epitaxial layer and reduce the width of the epitaxial layer. However, this reduces the reverse-biased voltage which the "off" transistor can withstand before the junction breaks down and increases the collector junction capacitance — two undesirable effects.

2. R'_C can be reduced by decreasing the sheet resistance of the base region underneath the emitter either by widening the base or increasing the impurity doping in the base. However, this will reduce the switching time of the transistor and limit the maximum frequency of the signal which can be transmitted through the "on" transistor.

3. The emitter-base perimeter can be increased. This will not degrade other switching properties of the transistor, but the maximum value of the emitter-base perimeter is limited somewhat by space considerations. However, this does point out the advantage of using an emitter configuration that is interleaved with the base contacts.

Therefore the design of an epitaxial transistor for switching applications in a radiation environment still remains a compromise between breakdown voltage, saturation voltage, collector capacitance, and switching time. The most practical compromise to make will depend upon the particular application of the transistor and which of the above mentioned properties are most critical.

This investigation has also brought out an important point in considering the d-c current levels at which to operate the saturated transistor. Before irradiation,

very little advantage can be gained by increasing the base current beyond that required to drive the transistor into saturation. In fact, it is desirable not to increase the base current past this value since this will increase the input voltage drop V_{BE}. However, if the transistor is to be exposed to fast-neutron irradiation, a significant decrease in $V_{CE(SAT)}$ can be gained by increasing the base current to several times the value required to drive the transistor into saturation. In addition to the consideration of the base current, the d-c collector and consequently the emitter current should be kept as low as possible to reduce the $I_C R'_C$ drop and to reduce the crowding of the current entering the collector.

The effects of reduced minority-carrier lifetime on the collector-emitter saturation voltage discussed in this thesis are also applicable to switching transistors that have gold doping in the base and collector regions. Gold doping is a common practice in the manufacture of switching transistors in order to reduce carrier lifetime, thus reducing switching time. $V_{CE(SAT)}$ increases with gold doping not only due to the reduction in carrier lifetime but also due to the fact that gold doping causes a greater increase in the resistivity of the semiconductor regions than fast-neutron irradiation levels less than 10^{14} neutrons-cm^{-2}. Thus, the design of gold-doped switching transistors with fast switching times involves many of the same problems and compromises as the design of switching transistors for use in radiation environments.

APPENDIX A
VALIDITY OF ASSUMPTIONS

Assumption (1) limits the geometry being considered to that most frequently employed in transistors, including the epitaxial transistors considered in this investigation. Assumption (2) can be shown to be a valid assumption on the basis of the resistivity and widths of the emitter region and collector substrate. Both of these regions have a resistivity of approximately 0.0001 ohm-cm. Assuming a maximum emitter width of 3×10^{-4} cm and minimum area of 10^{-3} cm^2, the resistance of the emitter area can be calculated as

$$R_E = \frac{\rho_e W_e}{A} = \frac{10^{-4} \times 3 \times 10^{-4}}{10^{-3}} = 3 \times 10^{-5}\ \Omega \qquad (1)$$

For the collector substrate, the maximum width is 5 mils or $\approx 13 \times 10^{-3}$ cm. Assuming as a minimum effective collector area the emitter area above, the resistance of the substrate can be calculated to be

$$R_{C_{SUB}} = \frac{10^{-4} \times 13 \times 10^{-3}}{10^{-3}} \approx 10^{-3}\ \Omega \qquad (2)$$

Thus, for maximum currents of $I_C = 3$ amperes and $I_E = 4.2$ amperes, the voltage drops in these regions will be negligible compared to the other voltages in the transistor.

Assumption (3) is made because at the relatively high current densities considered in this investigation, the maximum surface recombination current when the oxide-coated surface is saturated is negligible compared to the current densities in the base. Previous work on this type of transistor has shown that surface recombination current can be neglected when considering the operation of the transistor.

Assumption (4) is made in an effort to simplify the analysis as much as possible in cases for which the step junction approximation is reasonable. The effects of the impurity gradient in the base are included in the base sheet resistance and average resistivity. The values of these parameters used in the analyses were either measured on these transistors or obtained from the manufacturer.

The validity of Assumption (5) is not obvious from the data presented in this thesis; rather, it is based upon measurements of junction breakdown voltages and

capacitance made on these transistors before and after fast-neutron irradiation. These measurements indicate that the carrier concentrations are changed by no more than a few percent for fast-neutron flux levels up to 10^{14} neutrons/cm^2. Conductivity (σ) can be expressed in terms of the carrier concentrations (n, p) by

$$\sigma = q\left(\mu_n n + \mu_p p\right) \text{ ohm-cm} \qquad (3)$$

Thus the conductivity would be expected to change by no more than a few percent for the above mentioned fast-neutron flux level. It should be noted, however, that at fast-neutron flux levels above 10^{14} neutrons/cm^2, the change in conductivity may become significant and thus have to be considered in determining operation of semiconductor devices at high levels of fast-neutron irradiation.

APPENDIX B
EBERS AND MOLL RELATIONSHIP FOR JUNCTION VOLTAGE VERSUS JUNCTION CURRENT

The Ebers and Moll[17] relationship is based upon the assumption that current-voltage relations for the junction can be written in the form

$$I = I_S\left[\exp\left(\frac{qV}{kT}\right) - 1\right] \tag{1}$$

This assumption is valid only for junctions for which the minority-carrier concentration injected into a semiconductor region is negligible compared to the majority-carrier concentration in that region. However, in the operation of many silicon transistors, this assumption is not valid because the concentration of minority carriers injected into the base region or collector region may exceed the majority-carrier concentration by more than an order of magnitude. In this case, it has been shown[23] that the diode equation can be written as

$$I = I_S\left[\exp\left(\frac{qV}{nkT}\right) - 1\right] \tag{2}$$

where $1 \leq n \leq 2$ depending upon the relative concentrations of minority and majority carriers involved.

Thus the collector and emitter currents (assumed positive when flowing into transistor) can be written as

$$I_E = A_{11}\left[\exp\left(\frac{qV_E}{nkT}\right) - 1\right] + A_{12}\left[\exp\left(\frac{qV_C}{nkT}\right) - 1\right] \tag{3}$$

$$I_C = A_{21}\left[\exp\left(\frac{qV_E}{nkT}\right) - 1\right] + A_{22}\left[\exp\left(\frac{qV_C}{nkT}\right) - 1\right] \tag{4}$$

where V_E and V_C are the emitter and collector junction voltages respectively. Under normal transistor operation, the collector junction is reverse biased ($V_C << 0$); thus I_E and I_C can be written as

$$I_E = A_{11}\left[\exp\left(\frac{qV_E}{nkT}\right) - 1\right] - A_{12} \tag{5}$$

$$I_C = A_{21}\left[\exp\left(\frac{qV_E}{nkT}\right) - 1\right] - A_{22} \tag{6}$$

Solving Equation (5) for $[\exp(qV_E/nkT) - 1]$ and substituting this value in Equation (6) yields

$$I_C = \frac{A_{21}}{A_{11}} I_E + \left(\frac{A_{12}A_{21}}{A_{11}} - A_{22} \right) \tag{7}$$

Equation (7) can be written in terms of transistor parameters as

$$I_C = -\alpha_N I_E + I_{CO} \tag{8}$$

For inverse transistor operation, the emitter junction is reverse biased and a similar procedure as used above for I_C can be used to determine I_E as

$$I_E = \frac{A_{12}}{A_{22}} I_C + \left(\frac{A_{12}A_{21}}{A_{22}} - A_{11} \right) \tag{9}$$

For the inverse transistor, I_E can be written as

$$I_E = -\alpha_I I_C + I_{EO} \tag{10}$$

Comparing Equations (7), (8), (9), and (10), the A's can be evaluated as

$$A_{11} = -\frac{I_{EO}}{1 - \alpha_N \alpha_I} \tag{11}$$

$$A_{22} = -\frac{I_{CO}}{1 - \alpha_N \alpha_I} \tag{12}$$

$$A_{12} = \frac{\alpha_I I_{CO}}{1 - \alpha_N \alpha_I} \tag{13}$$

$$A_{21} = \frac{\alpha_N I_{EO}}{1 - \alpha_N \alpha_I} \tag{14}$$

Using the above equations, I_E and I_C can now be written as

$$I_E = -\frac{I_{EO}}{1 - \alpha_N \alpha_I} \left[\exp\left(\frac{qV_E}{nkT} \right) - 1 \right] + \frac{\alpha_I I_{CO}}{1 - \alpha_N \alpha_I} \left[\exp\left(\frac{qV_C}{nkT} \right) - 1 \right] \tag{15}$$

$$I_C = \frac{\alpha_N I_{EO}}{1 - \alpha_N \alpha_I} \left[\exp\left(\frac{qV_E}{nkT} \right) - 1 \right] - \frac{I_{CO}}{1 - \alpha_N \alpha_I} \left[\exp\left(\frac{qV_C}{nkT} \right) - 1 \right] \tag{16}$$

Equations (15) and (16) can be solved for the exponential factors and simplified to yield

$$-I_{EO}\left[\exp\left(\frac{qV_E}{nkT}\right) - 1\right] = I_E + \alpha_I I_C \tag{17}$$

$$-I_{CO}\left[\exp\left(\frac{qV_C}{nkT}\right) - 1\right] = \alpha_N I_E + I_C \tag{18}$$

Solving for V_E and V_C yields

$$V_E = \frac{nkT}{q}\,\ell n\left(-\frac{I_E + \alpha_I I_C}{I_{EO}} + 1\right) \tag{19}$$

$$V_C = \frac{nkT}{q}\,\ell n\left(-\frac{I_C + \alpha_N I_E}{I_{CO}} + 1\right) \tag{20}$$

For any reasonable values of I_E and I_C, Equations (19) and (20) can be reduced to

$$V_E = \frac{nkT}{q}\,\ell n\left(-\frac{I_E + \alpha_I I_C}{I_{EO}}\right) \tag{21}$$

$$V_C = \frac{nkT}{q}\,\ell n\left(-\frac{I_C + \alpha_N I_E}{I_{CO}}\right) \tag{22}$$

Using the facts that $I_E + I_C + I_B = 0$ and that $\alpha_I I_{CO} = \alpha_N I_{EO}$, the voltage drop across the junctions can be expressed in terms of base and collector current as

$$V_{CE} = \pm\left(V_C - V_E\right) \tag{23}$$

$$V_{CE} = \pm\frac{nkT}{q}\,\ell n\left\{\frac{\alpha_I\left[1 - I_C\left(1 - \alpha_N\right)/\left(I_B\alpha_N\right)\right]}{1 + I_C/I_B\left(1 - \alpha_I\right)}\right\} \tag{24}$$

The (+) sign applies to PNP transistors and the (-) sign to NPN transistors. Using Equation (24) with $n > 1$ is valid only if high-level injection occurs to the same degree at both the collector and emitter junctions. For example, if high-level injection occurs at the collector junction but not at the emitter junction, the following modification of Equation (24) would be more accurate:

$$V_{CE} = \pm\frac{kT}{q}\left[\ell n\left\{\frac{\alpha_I\left[1 - I_C\left(1 - \alpha_N\right)/\left(I_B\alpha_N\right)\right]}{1 + I_C/I_B\left(1 - \alpha_I\right)}\right\} + (n - 1)\,\ell n\left(-\frac{I_C + \alpha_N I_E}{I_{CO}}\right)\right] \tag{25}$$

For the transistors examined in this investigation and for the current levels considered, another complication arises from the fact that the current through the collector junction is composed of electron and hole currents. The hole current injected into the collector from the base is considered as high-level injection due to the low impurity concentration in the epitaxial layer. However, the electron current through the epitaxial layer is majority-carrier current, and high-level injection does not apply. Due to the fact that the hole current is much smaller than the electron current, it can be assumed that Equation (24) gives a good approximation of V_{CE} with $n \approx 1$.

APPENDIX C
CONDUCTIVITY MODULATION

Conductivity modulation results whenever minority carriers are injected into a semiconductor region in a concentration that is of the same order of magnitude or greater than the concentration of majority carriers present in the region. In order to maintain charge neutrality, the majority carriers must increase to the same concentration as that of the minority carriers, thus increasing the number of free carriers in the semiconductor and increasing the conductivity, or, in other words, reducing the resistivity. For example, consider an n-type semiconductor for which $n >> p$. The resistivity can be expressed as

$$\rho = \frac{1}{q\mu_n n + q\mu_p p} \approx \frac{1}{q\mu_n n} \ \Omega\text{-cm} \tag{1}$$

Since $n \approx N_D$, where N_D is the concentration of the donor atoms, the resistivity can be written as

$$\rho \approx \frac{1}{q\mu_n N_D} \tag{2}$$

If a hole current is injected into the n-region, the concentration of holes injected is given by

$$p \cdot \frac{W J_p}{2q D_p} \tag{3}$$

if the width of the n-region (W) is less than the diffusion length for holes (L_p) in the n-region. If L_p is less than W, then L_p should be used in Equation (3) and only a distance of L_p can be considered to have an excess hole concentration. Since n must equal p, the resistivity of the n-region is now given by

$$\rho' = \frac{1}{q\mu_n N_D \left(1 + W J_p / 2q D_p N_D\right)} \ \Omega\text{-cm} \tag{4}$$

or

$$\rho' = \frac{\rho}{\left(1 + W J_p / 2q D_p N_D\right)} \ \Omega\text{-cm} \tag{5}$$

The hole current density at which the effects of conductivity modulation becomes significant can be expressed from Equation (3) as

$$J_p = \frac{2q\, D_p N_D}{W} = \frac{2\, kT}{q}\,\frac{1}{W_p} \qquad (6)$$

APPENDIX D
INJECTION EFFICIENCY OF COLLECTOR

An approximate expression for the injection efficiency of the collector of an epitaxial transistor can be derived, based on the assumption that the N^+ substrate acts like a metal contact and reduces the hole concentration in the epitaxial region to its equilibrium value at the N^-N^+ junction. This is shown in Figure II-D-1.

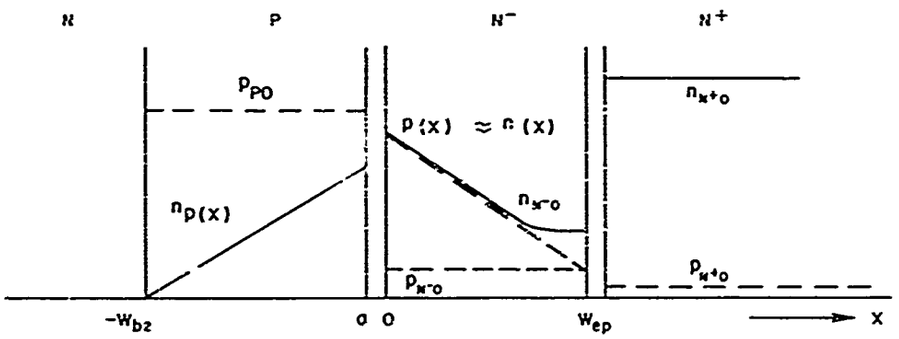


Figure II-D-1. Carrier Concentrations Which Determine Injection Efficiency of Collector

In the N^- region, the hole concentration can be expressed as

$$P(x) - P_{N^-0} = A \exp\left(\frac{x}{L_{pc}}\right) + B \exp\left(\frac{-x}{L_{pc}}\right) \tag{1}$$

To determine the coefficients A and B, Equation (1) must be evaluated at the boundaries x = 0 and $x = W_{ep}$.

At x = 0

$$P = P_{N^-0} \exp\left(\frac{qV_C}{kT}\right) \tag{2}$$

and at $x = W_{ep}$

$$P = P_{N^-0} \tag{3}$$

where V_C is the forward bias voltage across the collector junction (PN^-).

Substituting Equations (2) and (3) into (1), A and B can be evaluated as

$$A = \frac{P_{N^-0}\left[1 - \exp\left(qV_C/kT\right)\right]\exp\left(-W_{ep}/L_{pc}\right)}{2\sinh\left(W_{ep}/L_{pc}\right)} \tag{4}$$

$$B = \frac{-P_{N^-0}\left[1 - \exp\left(qV_C/kT\right)\right]\exp\left(W_{ep}/L_{pc}\right)}{2\sinh\left(W_{ep}/L_{pc}\right)} \tag{5}$$

where W_{ep} is the width of the epitaxial region. Substituting for A and B in Equation (1) and simplifying yields

$$P(x) = P_{N^-0} - P_{N^-0}\left[1 - \exp\left(\frac{qV_C}{kT}\right)\right]\frac{\sinh\left[\left(W_{ep} - x\right)/L_{pc}\right]}{\sinh\left(W_{ep}/L_{pc}\right)} \tag{6}$$

The hole current density injected into the N^- epitaxial region from the P base can be expressed as

$$J_p = q\,D_{pc}\left.\frac{dp(x)}{dx}\right|_{x=0} \tag{7}$$

Performing the differentiation and setting x = 0 yields

$$J_p = \frac{q\,D_{pc}\,P_{N^-0}}{L_{pc}}\left[\exp\left(\frac{qV_C}{kT}\right) - 1\right]\frac{1}{\tanh\left(W_{ep}/L_{pc}\right)} \tag{8}$$

The electron current density injected from the epitaxial region into the base can be determined from the electron concentration at x = a by

$$J_n = \frac{q\,D_{nb}\,n(a)}{W_{b2}} \tag{9}$$

Using the fact that n(0) = P(0), then n(a) can be expressed as

$$n(a) = P(0)\exp\left(\frac{qV_C}{kT}\right)\exp\left(\frac{-qV_0}{kT}\right) \tag{10}$$

But

$$\exp\left(\frac{-qV_0}{kT}\right) = \left(\frac{P_{N^-0}}{P_{po}}\right) \tag{11}$$

and

$$P(0) = P_{N^-0}\exp\left(\frac{qV_C}{kT}\right) \tag{12}$$

Therefore n(a) can be written as

$$n(a) = \frac{P_{N^-0}^2}{P_{po}} \exp\left(\frac{2qV_C}{kT}\right) \tag{13}$$

and J_n can be written as

$$J_n = \frac{q\, D_{nb}}{W_{b2}} \frac{P_{N^-0}^2}{P_{po}} \exp\left(\frac{2qV_C}{kT}\right) \tag{14}$$

The injection efficiency of the collector is defined as

$$\gamma_I = \frac{J_n}{J_n + J_p} = 1 + \left(\frac{J_p}{J_n}\right)^{-1} \tag{15}$$

Using the approximation that $\exp\left(qV_C/kT\right) >> 1$, the ratio J_p/J_n can be written as

$$\frac{J_p}{J_n} = \frac{D_{pc}\, W_{b2}\, P_{po} \exp\left(-qV_C/kT\right)}{D_{nb}\, L_{pc}\, P_{N^-0} \tanh\left(W_{ep}/L_{pc}\right)} \tag{16}$$

Equation (14) can be written in terms of J_p using Equation (8) as

$$\frac{J_p}{J_n} = \frac{q\, D_{pc}^2\, W_{b2}\, P_{po}}{D_{nb}\, L_{pc}^2\, J_p \tanh^2\left(W_{ep}/L_{pc}\right)} \tag{17}$$

Thus it can be seen that J_p/J_n is a function of J_p and also L_{pc}, the diffusion length for holes in the epitaxial layer.

If $L_{pc} >> W_{ep}$, $\tanh\left(W_{ep}/L_{pc}\right) \approx W_{ep}/L_{pc}$ and

$$\frac{J_p}{J_n} = \frac{q\, D_{pc}^2\, W_{b2}\, P_{po}}{D_{nb}\, W_{ep}^2\, J_p} \tag{18}$$

If $W_{ep} > 2\, L_{pc}$, $\tanh\left(W_{ep}/L_{pc}\right) \approx 1$ and

$$\frac{J_p}{J_n} = \frac{q\, D_{pc}^2\, W_{b2}\, P_{po}}{D_{nb}\, L_{pc}^2\, J_p} \tag{19}$$

A numerical example will give an indication of the magnitude of γ_I. Assuming a reasonable value for J_p of 100 amp/cm^2, γ_I can be calculated for the Clevite 3TX002 before irradiation using

$D_{pc} \approx 10\ cm^2/sec$

$D_{nb} \approx 10\ cm^2/sec$

$W_{ep} = 2 \times 10^{-3}\ cm << L_p$

$P_{po} = 5 \times 10^{-17}\ cm^{-3}$

$W_{b2} = 5 \times 10^{-4}\ cm$

$\frac{J_p}{J_n} = 1$ and $\gamma_I = 0.5$

APPENDIX E
LIST OF SYMBOLS

A	Area	N_D	Donor impurity concentration
A'	Effective area	n	Electron concentration
α_N	Grounded-base current gain of normal transistor	p	Hole concentration
α_I	Grounded-base current gain of inverted transistor	P_{po}	Equilibrium hole concentration in P region
BV_{CBO}	Collector-base breakdown voltage	P_{NO}	Equilibrium hole concentration in collector epitaxial layer
$\beta_N = h_{FE}$	Grounded-emitter current gain of normal transistor	q	Charge of electron
β_N^*	Base-transport factor of normal transistor	R_B	Base resistance
β_I^*	Base-transport factor of inverted transistor	R'_B	Base resistance between normal and inverted transistors
C_{TC}	Collector junction capacitance	R'_C	Effective resistance of epitaxial layer
γ_N	Emitter efficiency of normal transistor	R_D	Lateral resistance in epitaxial region
γ_I	Emitter efficiency of inverted transistor	ρ_{b1}	Resistivity of base region under emitter
D_{nb}	Electron diffusion constant in base	ρ_{b2}	Resistivity of base region not under emitter
D_{pc}	Hole diffusion constant in collector epitaxial layer	ρ_{ep}	Resistivity of epitaxial region
f_T	Cutoff frequency	ρ_e	Resistivity of emitter
h	Emitter-base perimeter	σ	Conductivity
I_B	Total base current	T	Absolute temperature
I_C	Total collector current	τ_{nb}	Electron lifetime in base
I_{FC}	Collector forward current	τ_{nbi}	Electron lifetime in base before irradiation
I_{RC}	Collector reverse current	τ_{pc}	Hole lifetime in collector epitaxial layer
I_E	Total emitter current	V_{JC}	Collector junction voltage
I_{FE}	Emitter forward current	V_{JE}	Emitter junction voltage
I_{RE}	Emitter reverse current	V_{CE}	Collector-emitter voltage
I_{CO}	Saturation current of collector junction with zero emitter current	V_{BE}	Emitter-base voltage
I_{EO}	Saturation current of emitter junction with zero collector current	V_C	Collector junction bias voltage
J_p	Hole current density	V_O	Junction contact potential
k	Boltzmann's constant	V_R	Lateral voltage drop in base
L_{pc}	Hole diffusion length in collector epitaxial layer	V_D	Lateral voltage drop in collector
L_{pe}	Hole diffusion length in emitter	W_{b1}	Width of base underneath emitter
μ_n	Electron mobility	W_{b2}	Width of base not underneath emitter
μ_p	Hole mobility	W_{ep}	Width of epitaxial region
		ϕ	Integrated fast-neutron flux level

LIST OF SYMBOLS*

A	Area
A'	Effective area
α_N	Grounded-base current gain of normal transistor
α_I	Grounded-base current gain of inverted transistor
BV_{CBO}	Collector-base breakdown voltage
$\beta_N = h_{FE}$	Grounded-emitter current gain of normal transistor
β_N^*	Base-transport factor of normal transistor
β_I^*	Base-transport factor of inverted transistor
C_{TC}	Collector junction capacitance
γ_N	Emitter efficiency of normal transistor
γ_I	Emitter efficiency of inverted transistor
D_{nb}	Electron diffusion constant in base
D_{pc}	Hole diffusion constant in collector epitaxial layer
f_T	Cutoff frequency
h	Emitter-base perimeter
I_B	Total base current
I_C	Total collector current
I_{FC}	Collector forward current
I_{RC}	Collector reverse current
I_E	Total emitter current
I_{FE}	Emitter forward current
I_{RE}	Emitter reverse current
I_{CO}	Saturation current of collector junction with zero emitter current
I_{EO}	Saturation current of emitter junction with zero collector current
J_p	Hole current density
k	Boltzmann's constant
L_{pc}	Hole diffusion length in collector epitaxial layer
L_{pe}	Hole diffusion length in emitter
μ_n	Electron mobility
μ_p	Hole mobility
N_D	Donor impurity concentration
n	Electron concentration
p	Hole concentration
P_{po}	Equilibrium hole concentration in P region
P_{NO}	Equilibrium hole concentration in collector epitaxial layer
q	Charge of electron
R_B	Base resistance
R'_B	Base resistance between normal and inverted transistors
R'_C	Effective resistance of epitaxial layer
R_D	Lateral resistance in epitaxial region
ρ_{b1}	Resistivity of base region under emitter
ρ_{b2}	Resistivity of base region not under emitter
ρ_{ep}	Resistivity of epitaxial region
ρ_e	Resistivity of emitter
σ	Conductivity
T	Absolute temperature
τ_{nb}	Electron lifetime in base
τ_{nbi}	Electron lifetime in base before irradiation
τ_{pc}	Hole lifetime in collector epitaxial layer
V_{JC}	Collector junction voltage
V_{JE}	Emitter junction voltage
V_{CE}	Collector-emitter voltage
V_{BE}	Emitter-base voltage
V_C	Collector junction bias voltage
V_O	Junction contact potential
V_R	Lateral voltage drop in base
V_D	Lateral voltage drop in collector
W_{b1}	Width of base underneath emitter
W_{b2}	Width of base not underneath emitter
W_{ep}	Width of epitaxial region
ϕ	Integrated fast-neutron flux level

*Appendix E repeated.

REFERENCES

1. J. H. Crawford and J. W. Cleland, "Radiation Effects in Semiconductors," Progress in Semiconductors, ed. A. F. Gibson, Vol. 2 (New York: John Wiley and Sons, 1957).
2. D. S. Billington and J. H. Crawford, Radiation Damage in Solids (Princeton, New Jersey: Princeton University Press, 1961) 312.
3. G. J. Dienes and G. H. Vineyard, Radiation Effects in Solids, Monographs in Physics and Astronomy, Vol. II (New York: Interscience Publishers, 1957).
4. H. Y. Fan and K. Lark-Horovitz, "Effects of Radiation on Semiconductors," The Effects of Radiation on Materials, ed. J. J. Harwood et al. (New York: Reinhold Publishing Corp., 1958).
5. K. Lark-Horovitz, "Nucleon-bombarded Semiconductors," Semiconducting Materials, ed. H. K. Henisch (New York: Academic Press, 1951) 47.
6. F. Seitz and J. S. Koehler, "Displacement of Atoms during Irradiation," Solid State Physics, Vol. 2 (New York: Academic Press, 1956) 305.
7. E. E. Klontz and K. Lark-Horovitz, "Electron Bombardment of Ge," Phys. Rev. Vol. 82 (1951) 763.
8. H. M. James and K. Lark-Horovitz, "Localized Electronic States in Bombarded Semiconductors," Z. physik. Chem., Vol. 198 (1951) 107.
9. J. W. Cleland et al., "The Effect of Fast Neutron Bombardment on the Electrical Properties of Germanium," Phys. Rev., Vol. 83 (1951) 318.
10. G. Bemski, "Recombination in Semiconductors," Proc. IRE, Vol. 46 (June 1958) 990.
11. R. W. Beck, E. Paskell, and C. S. Peet, "Some Effects of Fast Neutron Irradiation on Carrier Lifetimes in Silicon," J. Appl. Phys., Vol. 30, No. 9 (1959) 1437.
12. J. R. Bilinski and R. Merrill, "Selecting Transistors for Radiation Environments," Electronics, Vol. 32 (Dec. 25, 1959) 40.
13. J. J. Loferski, "Analysis of the Effect of Nuclear Radiation on Transistors," J. Appl. Phys., Vol. 29, No. 1 (1958) 35.
14. I. C. Verster, "Silicon Planar Epitaxial Transistors as Fast and Reliable Low-Level Switches," IEEE Trans. on Electron Devices, Vol. ED-11 (May 1964) 228.

15. A. B. Phillips, Transistor Engineering (New York: McGraw-Hill Book Co., 1962) Chapters 1 and 17.

16. L. P. Hunter, Handbook of Semiconductor Electronics (New York: McGraw-Hill Book Co., 1956) 4-14.

17. J. J. Ebers and J. L. Moll, "Large-Signal Behavior of Junction Transistors," Proc. IRE, Vol. 42 (Dec. 1954) 1761.

18. J. L. Moll, "Large-Signal Transient Response of Junction Transitors," Proc. IRE, Vol. 42 (Dec. 1954) 1773.

19. H. G. Rudenberg, "On the Effect of Base Resistance and Collector-to-Base Overlap on the Saturation Voltages of Power Transistors," Proc. IRE, Vol. 46 (June 1948) 1304.

20. J. R. Hauser, "The Effects of Distributed Base Potential on Emitter-Current Injection Density and Effective Base Resistance for Stripe Transistor Geometries," IEEE Trans. on Electron Devices, Vol. ED-11 (May 1964) 238.

21. N. J. Chaplin, "Injected Carrier Concentrations in Diffused Silicon Transistors by Observations of Interband Carrier Recombination," presented at the Professional Group on Electron Devices Meeting (IRE), Washington, D. C., October 26, 1961.

22. J. A. Champion, "Potential Distribution in a Semiconductor Block with Strip Contacts," Solid State Electronics, Vol. 8 (April 1965) 431.

23. A. B. Phillips, Transistor Engineering (New York: McGraw-Hill Book Co., 1962) Chapter 6.

ATTACHMENT III

PERMANENT RADIATION DAMAGE EFFECTS IN NARROW BASE PNPN DEVICES

1. INTRODUCTION

The rapid degradation in performance of wide-base PNPN devices by nuclear radiation has been pointed out by a number of investigators.[1-3] As a result of these studies a stigma of poor radiation performance has come to be associated with four-layer structures. This is not a valid conclusion. In this paper the effects of fast-neutron irradiation on narrow-base PNPN triodes will be discussed, and results will show that, for comparable base widths, PNPN devices are far more resistant to radiation damage than are bipolar transistors. This relative insensitivity to radiation damage is due to the collinear flow of majority and minority carriers in the bases of a PNPN, in contrast to the case of bipolar transistors where majority carrier flow in the base is perpendicular to minority carrier flow. The collinear flow of carriers in the PNPN gives rise to field-aided transport in both base regions. As a result, the minority carriers are caused to move across the bases more rapidly than in the case of diffusion-controlled bipolar devices. Thus, for comparable base widths, the PNPN can tolerate a far shorter lifetime than can the transistor Results will also show that, in lieu of increasing the radiation tolerance of the PNPN device over the ordinary transistor, one can at the same level of radiation dose achieve much higher power handling capabilities in the PNPN structure. Results obtained with currently available PNPN devices indicate more than an order of magnitude improvement in power handling capability over bipolar transistors.

The major effects of nuclear irradiation of semiconductor devices are due to the degradation of the excess carrier lifetimes and, to a lesser extent, to the changes in the majority carrier concentration. In a minority carrier device where the operation depends on injection, decay, and transport of excess carrier densities, the effects of a short lifetime may be better visualized in terms of diffusion length in relation to device geometry, i.e., the basewidth to diffusion length ratio W/L. Analyses in terms of this parameter as applied to $P^+N^-N^+$ structures can be found in several sources. Stafeev[4] stresses the properties of diodes with a large ratio and demonstrates the strong dependence of the forward current on the diffusion length. Mayer, Baron, and Marsh[5] show that in long silicon

PIN structures, diffusion significantly affects the current voltage characteristics even where the W/L ratio is much greater than ten.

The $P^+N^-PN^+$ device can be thought of as a $P^+N^-N^+$ structure with an additional center junction which provides a blocking characteristic in the forward direction up to some characteristic current, the holding current. Above this current the lightly doped N^- base is inundated with holes from the P^+ region and electrons from the N^+PN^- transistor. The properties of such $P^+N^-PN^+$ devices with large W/L ratios have been discussed by Aldrich and Holonyak[6] and Kuzmin.[7] Generally these $P^+N^-PN^+$ devices have a relatively wide N^- base, since the width of this base determines the blocking voltage capability. For a maximum blocking voltage capability this base will not only be relatively wide but of fairly high-resistivity material. In the typical manufacturing process the starting material of the silicon wafer would be high-resistivity N type, which will later form the wide base region of the PNPN. A deep P type emitter diffusion is done to form one emitter junction, and a typical transistor base and emitter diffusion is carried out on the other side of the wafer to form the remainder of the structure. Thus, the narrower base transistor will usually be the NPN transistor with a basewidth typically 2 - 5 μ. This technique results in bases for the PNP transistor typically 50 - 200 μ wide. As a consequence of the wide N base and the large W/L ratio, these devices are rather sensitive to permanent damage effects of fast neutrons.

In this paper, permanent damage effects in narrow N^- base (15 - 20μ) $P^+N^-PN^+$ structures will be discussed. These devices are made by epitaxial deposition of the N^- base on a heavily doped P^+ substrate. Ordinary transistor diffusion processes are then used to form the remaining P base and N^+ emitter. Such devices have a relatively low blocking voltage capability but a high resistance to radiation damage. In comparison with bipolar transistors of comparable basewidths they are significantly less sensitive to damage and for the same radiation exposure have a much greater power switching capability.

The superiority of the PNPN over the bipolar transistor is a result of the fact that majority carrier currents to both base regions of the PNPN flow parallel rather than perpendicular to the minority carrier flow. Consequently the current crowding effect of ordinary transistors is absent. In addition, the PNPN majority carrier base current gives rise to an electric field in the lightly doped N base which aids the minority carrier transport across the base regions. Thus PNPN switching ($\alpha_1 + \alpha_2 \geq 1$) will occur and sustain the "on" condition even though the diffusion length of the minority carriers in the N base region is small compared to its basewidth, i.e., $L_p << W_n$. Ordinary transistors would require that $L_p >> W_n$ for satisfactory operation.

In this paper, the behavior of PNPN devices and the transport of minority carriers in the two base regions will be discussed. Then, the experimental results obtained on a variety of PNPN structures subjected to fast neutron environments will be presented. Following these results, the design of radiation-hardened devices and the relative power switching capabilities of PNPN's and transistors will be considered.

2. PHENOMENOLOGICAL EXPLANATION OF A PNPN DEVICE AND COMPARISON WITH BIPOLAR TRANSISTOR

a. PNPN Operation

A PNPN device may be thought of as a device with inherent PNP and NPN transistors forming a regenerative feedback pair.*[8,9] The base current or the recombination current and nonunity emitter efficiency current at the emitter junction are supplied mutually and collinearly by the collector current of the other transistor. Thus, as the loop gain approaches unity, each transistor drives its mate into saturation. By balancing the constituent currents at the center junction and introducing the current amplification factors (α_1, α_2) of the minority carriers injected respectively by the exterior P^+ region and the N^+ region, one readily finds three regions in the V-I characteristics: (1) a forward blocking state corresponding to $\alpha_1 + \alpha_2 < 1$; (2) a negative resistance transistor region corresponding to $\alpha_1 + \alpha_2 \cong 1$; and (3) a forward on-state or the high-current, low-voltage region corresponding to $\alpha_1 + \alpha_2 \geq 1$. The transition from forward blocking to "on" states is normally facilitated by temporarily injecting additional majority carriers (gate current) into one of the base regions, usually the narrow P base, thus increasing the collector current which, in turn, increases the alphas and satisfies the "on" state requirement. However, once the device is driven into the "on" condition, all three junctions are forward biased and the gate drive is no longer necessary to sustain the "on" condition.

b. Comparison of PNPN and Bipolar Transistor

In order to compare the carrier flow in a PNPN and a transistor, schematic representations of both are given in Figure III-1. Here the flow of the majority and minority carriers in a PNPN device in the "on" condition is shown to be collinear, in contrast to transistors where the continuous base drive or the majority carrier flow is perpendicular to the minority carrier flow. In the transistor the resultant

---

*From the circuit point of view, the division of the thyristor into two coupled transistors may still be kept up formally, but from a device physics point of view, this concept is no longer meaningful at high injection levels and at short carrier lifetimes.

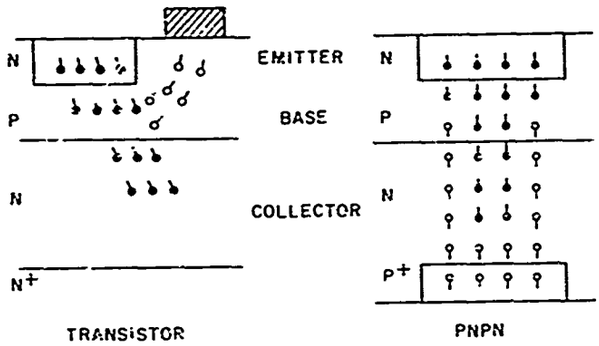


Figure III-1. Comparison of Carrier Flow in a Saturated Transistor and PNPN Devices in the "On" Condition

lateral base voltage drop causes current crowding at the emitter periphery and a localized current entering the collector region. If the transistor alpha is degraded by radiation damage, the base current increases, thus further intensifying the crowding effect. In the PNPN the collinear flow and the uniform current density along the junction are not altered even under irradiation. Thus there is considerable flexibility in designing very large emitter-area PNPN devices which are capable of carrying larger currents than transistors.

The requirement ($\alpha_1 + \alpha_2 \geq 1$) on the alphas for a PNPN is less than the requirement for a transistor ($\alpha \approx 1$). Hence, for the same diffusion length, wider bases in the PNPN are allowable and consequently higher punch-through and forward-blocking voltages can be achieved.

The collinear flow of majority and minority carriers produces a further advantage which is particularly evident in the wide N base of a PNPN in the "on" condition. The flow of majority carriers gives rise to an electric field which aids the flow of minority carriers across this region. By way of contrast, in the epitaxial layer of the collector region of a transistor, which must be comparable in width to the N base of the PNPN in order to obtain equivalent voltage switching capability, both minority and majority carriers enter from the base side. As a result the field due to the majority carrier flow opposes the flow of minority carriers. Thus for comparable W/L ratio the wide base of the PNPN will be more heavily conductivity-modulated than will be the collector region of the transistor. Since the voltage drops across these regions usually limit the power handling capabilities of irradiated PNPN's or bipolar transistors, the PNPN is clearly superior to the transistor in this respect. In Section 4 the question of the voltage drop across the wide base will be discussed in greater detail.

3. DISCUSSION OF MINORITY CARRIER TRANSPORT ALPHAS IN PNPN INCLUDING DRIFT EFFECTS

The PNPN devices investigated had base-width-to-diffusion-length ratios after irradiation of the order of unity in the narrow base region and much greater than unity in the wide base region. As pointed out previously, in such PNPN structures the operation of the device depends partially upon the electric field transport of minority carriers. The following paragraphs discuss the variation of the minority carrier current amplification factor with respect to changes in the base-width-to-diffusion-length ratio and the current injection level.

a. Blocking Region

At very low currents, in the order of 10^{-3} amp/cm^2 or less, both α_1 and α_2 increase with increasing current due to increasing emitter efficiencies and increasing transport factors. Since the N^- base is less doped than the P base, the variation of the voltage across the center junction and the resultant changes in the width of the space charge region cause appreciable changes in both the effective basewidth of the N^- region and its corresponding transport factor. At higher current ranges and in the negative resistance region, the voltage bias across the center junction goes through a transition from reverse through zero and then to a slightly forward condition. The rate of increase in alphas in this region is governed by the interaction of increased emitter efficiency, effective N^- basewidth, and the field-aided transport in the base region. At the end of the negative resistance or the holding condition, the center junction is slightly forward biased.

A simple analytical expression for the alpha of the lightly doped wide base region can be obtained if the following assumptions are made.

1. The forward bias at the center junction is negligible ($V_2 \ll kT/q$) so that the excess minority carrier concentration is zero at the center junction.
2. The injection level is low.
3. The electric field is approximately constant in the lightly doped wide base region.

The resulting calculation is similar to that given by Aldrich and Holonyak[6] and is based on the drift-enhanced diffusion model.

$$\alpha_1 = \frac{\left(\gamma_1 \sqrt{1 + V_1^2}\, e^{GV_1}\right) \Big/ \sinh\left(G\sqrt{1 + V_1^2}\right)}{V_1 + \sqrt{1 + V_1^2} \coth\left(G\sqrt{1 + V_1^2}\right)} \qquad (1)$$

where

$$V_1 = \frac{E_1 L_p}{2\,(kT/q)}, \quad G = \frac{W_n}{L_p}$$

E_1 is the electric field
L_p is the hole diffusion length
W_n is the width of the N^- base region
γ_1 is the emitter efficiency at the P^+N^- junction
k is Boltzmann's constant
T is temperature in degrees Kelvin
q is the electronic charge.

Figure III-2 shows a plot of the wide base alpha with respect to the voltage drop in half a diffusion length normalized by kT/q, indicating the strong dependence of alpha on electric field. Alpha based on a pure drift model, $\alpha_1 = \exp(-G/2V_1)$, is also plotted for comparison. For large basewidth-to-diffusion-length ratios, drift transport clearly predominates and large transport factors are possible even though W_n/L_p is large. The minority carriers are swept along by the electric field — a mechanism which, combined with the multiplication process at the center junction,

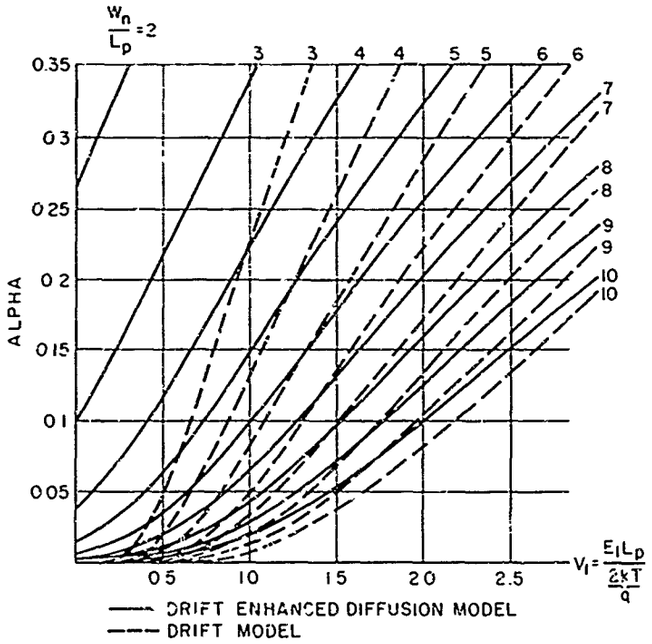


Figure III-2. Switching Characteristics of the Wide Base α in a PNPN Structure

enables the PNPN device to reach the turn-on condition even though the minority carrier lifetime is severely degraded.

In the narrow base region where W_n/L_p is approximately unity, the diffusion model, independent of current, is probably adequate for determining alpha in the vicinity of the turn-on process.

b. "On" Region

In the forward "on" condition, where high injection level conditions are assumed, there are two domains of interest. In the first the minority carrier density (hole concentration) in the wide base region increases linearly with the total current density J. Thus the electric field, and also the voltage drop across the wide base, is approximately independent of the current density. Alpha, which is roughly proportional to the product of the field intensity at the edge of the center junction and the ratio of the hole concentration at this edge to the total current density, is therefore approximately constant.

In the second domain, the injected minority hole density is proportional to the square of the current density which gives rise to the same relation for the base voltage drop. The average electric field is now proportional to the square root of the current density and the alpha is again approximately independent of the total current density.

Because of volume recombination and nonunity junction efficiencies, majority carrier current is caused to flow in the base regions, developing electric fields in both bases. One may also anticipate an asymmetrical field distribution in the lightly doped wide base region because the minority carrier concentration is lower at the N^-P center junction than at the P^+N^- junction, resulting in higher fields near the N^-P center junction. These electric fields, together with the forward bias at the center junction, cause the alphas of the two inherent transistors to be closely interrelated. This interdependence can be expressed in terms of the basewidth ratio, diffusion length ratio, doping ratio of the bases, and the emitter efficiencies at the exterior boundaries. When the lifetime of the minority carriers is degraded, the turn-on condition is satisfied at higher current levels through the aiding field in the lightly doped wide base region.

It is interesting to note the limiting values of alpha in the wide base region. At high injection level, it can be shown for large W/L ratios that if the wide base is N type, alpha cannot be greater than $1/(b + 1)$, and on the other hand if the wide base is P type, alpha will approach $b/(b + 1)$. For semiconductors of large mobility ratio b, one may anticipate a significant difference between P-type wide base and N-type wide base devices.

4. VOLTAGE DROP IN WIDE BASE

The V-I characteristics in the forward "on" condition of a PNPN device have been analyzed by Kuzmin.[7] With the simplifying assumptions that: (1) lifetime is independent of injection level, and (2) high level injection condition exists in the base region, the base voltage drop in the lightly doped wide base region is found to be

$$V_{N\ base} \cong \delta \left(\frac{W_n}{L_p}, \frac{W_p}{L_n}, \gamma_1, \gamma_3, \frac{N_A}{N_D}, b \right) \frac{kT}{q} e^{\left(W_n / 2\mathfrak{L}_p \right)} \tag{2}$$

where $\mathfrak{L}_p = L_p \sqrt{2b/(b+1)}$ is an elongated effective diffusion length due to the electric field. δ is a function of basewidth-to-diffusion-length ratio, the emitter efficiencies γ_1, γ_3 at P^+N and PN^+ junctions, the base doping ratio N_A/N_D, and the mobility ratio b. If the fall-off in the emitter efficiencies with increasing high injection current is neglected, δ takes a value between 1 and $2\pi/(b+1)$. The above expression indicates that the voltage drop in the wide base region is a sensitive exponential function of W_n/L_p and becomes the major contributor to the total voltage drop across the device as the lifetime of the minority carriers is decreased.

The algebraic sum of the three junction voltages is approximately equal to that of the P^+N junction, since the sum of the second and third junction is proportional to the logarithm of the ratio of the electron concentrations at the forward biased edges of the narrow base region, and is negligible. The sum of the junction voltages is then

$$V_J \cong \frac{kT}{q} \ln \frac{J\mathfrak{L}_p \left\{ \left[\gamma_1 (b+1) - 1 \right] \cosh W_n/\mathfrak{L}_p + \left[1 - (b+1)\alpha_1 \right] \right\}}{2qD_p P_n\, b \sinh W_n/\mathfrak{L}_p}$$

$$\cong \frac{kT}{q} \ln \frac{J\mathfrak{L}_p}{2qD_p P_n} \quad \text{for} \quad \frac{W_n}{\mathfrak{L}_p} >> 1 \text{ and } \gamma_1 = 1 \tag{3}$$

which indicates that the junction voltage is a logarithmic function of diffusion length and decreases slightly with decreasing minority carrier lifetime.

The condition at which the contribution from the wide base voltage drop begins to exceed that of junction voltage may be estimated from equations (2) and (3) as

$$J = \frac{2qD_p P_n}{\mathfrak{L}_p} e^{\left[\delta \exp \left(W_n/\mathfrak{L}_p \right) \right]} \tag{4}$$

Thus, for a given current level, devices with narrower base width, W_n, will have a much wider range of insensitivity to radiation damage. Figure III-3 shows the variations of the above expressions for the base voltage, and also the sum of the base

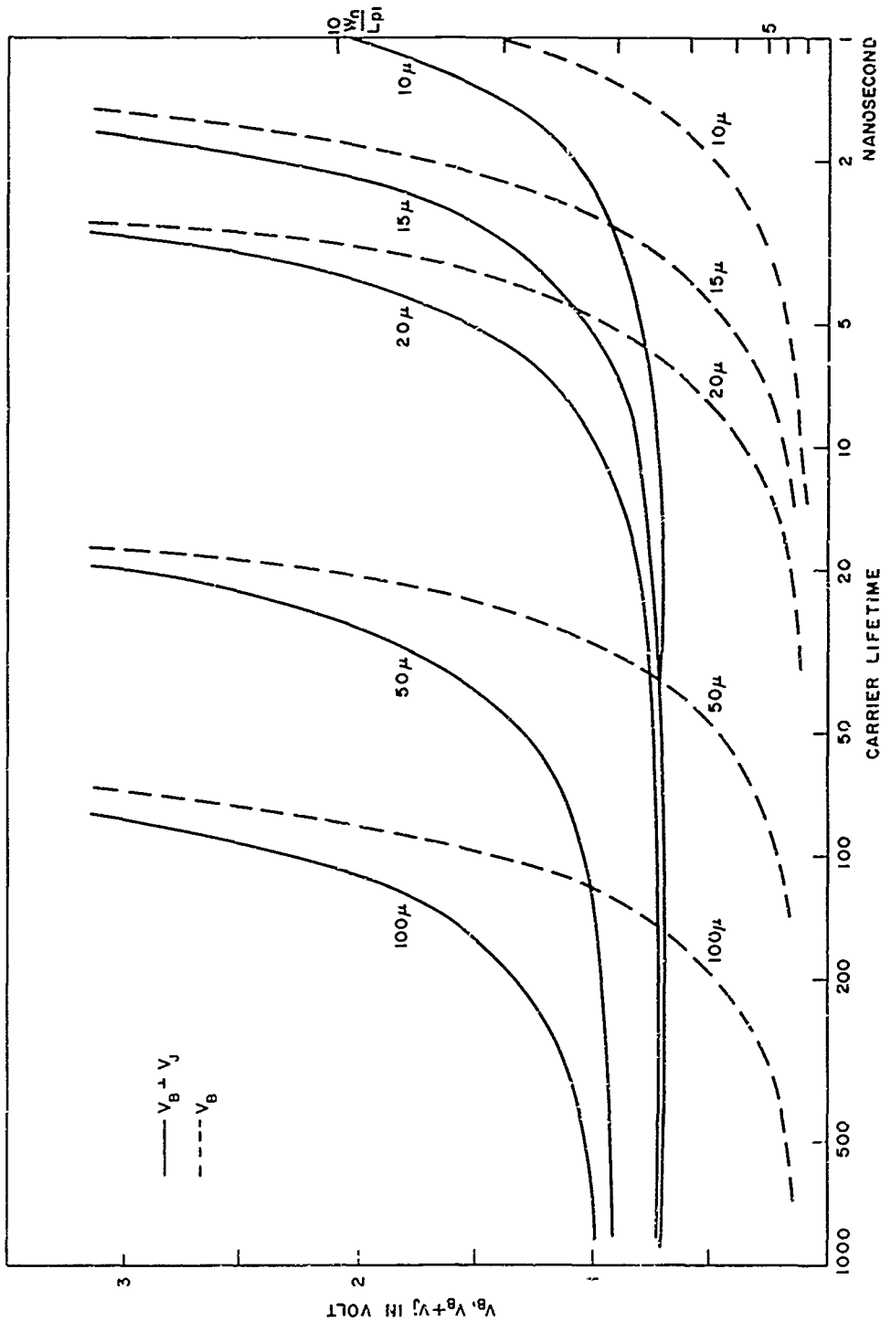


Figure III-3. Base and Junction Voltages versus Carrier Lifetime (I = 500 ma) Theoretical Value

voltage and the junction voltage plotted with respect to the minority carrier lifetime for devices with varying W/L.

5. EXPERIMENTAL RESULTS ON DEVICE PARAMETERS IN A NEUTRON ENVIRONMENT

a. Devices and Exposures

The dimensional parameters of the PNPN devices used in this investigation are shown in Table III-1.* Several types of conventional triple-diffused structures were also included in this study for comparison purposes. The first three types have relatively small emitter areas, while the rest of the devices have larger emitter areas and substantial gate-cathode interdigitation for decreasing the lateral spread time in the turn-on process. The four-lead devices were included in order to measure the alphas of the inherent transistors in the devices. The devices were exposed to fission spectrum neutrons at the Triga Reactor at Harry Diamond Laboratories and at the White Sands Fast Burst Reactor. Neutron fluences (E > 10 KeV) were determined with sulfur foil activation methods and are accurate to ≈ 20%. Pre- and post-irradiation measurements were made of the following parameters: reverse saturation current, alpha of the NPN part, hold and gate currents, "on" voltage, junction capacitance, breakdown voltage, and gate characteristics.

Table III-1

PNPN STRUCTURES

Device		W_p (μ)	W_n (μ)	ρ(N) (Ω-cm)	Emitter Area $cm^2 \times 10^{-3}$
ZB1002	(SSPI)	3	15	0.6	0.60
2N3032	(SSPI)	3	20	2.0	0.75
2N3032 (4L)*†	(SSPI)	3	25	8.0	0.75
ZB1001 (NB)†	(SSPI)	3	15	8.0	3.2
ZB1001 (WB)†	(SSPI)	3	20	8.0	3.2
3S2001 (WB)*	(SSPI)	6	50	25.0	3.7
2N1765 (WE)*	(WECo)	3-5	~100	25.0	4.5

*Triple-diffused structures.

†Four-lead devices.

*Solid State Products Inc., Salem, Massachusetts, provided most of the devices used in this study.

b. Reverse Saturation Current

In the forward blocking state, both exterior PN junctions are forward biased. The center junction is reverse biased and plays the role of a collector for the minority carriers injected into it by the exterior junctions. The total current at the reverse biased junction consists of electron and hole diffusion currents that originate at the exterior emitters and the generation current originating in the depletion region of the center junction.

At a very low current level (10^{-9} - 10^{-7} amp) the generation current at the center junction predominates over the diffusion components, which are minimized by the small transport factors. Thus a generation current I proportional to the recombination rate ($1/\tau$) is expected for the center junction. Figure III-4 shows the observed linear relation between the reverse saturation current and the neutron fluence (Φ) indicating that the minority carrier lifetime in the space charge region is linearly related to the reciprocal of the neutron fluence, i.e., $\tau \propto 1/\Phi$.

c. The Amplification Factor α_{NPN} for the Inherent NPN Transistor

It is difficult to measure the alpha of the inherent transistors of a PNPN under actual operating conditions, since any alpha measurement made with leads to the bases introduces external perturbations to the current flow and the voltage distribution.[10,11] However, the narrow base region generally has a basewidth-to-diffusion length of unity or less and a higher doping than the wide base, so the diffusion model may be adequate in describing the carrier flow in this region and the alpha measurement of this region is meaningful.

The calculated degradation of the alpha of the narrow base region with increasing neutron fluence assuming $1/\tau_{base} = 1/\tau_{no} + \Phi/K_n$ is shown in Figure III-5 with the corresponding experimental data taken on the four terminal devices. The measured values and the calculated values based on the usual transistor formula (α = sech W_p/L_n, assuming the emitter efficiency is nearly unity) agree for the following chosen parameters: the damage constant, $K_n = 10^6$ N/cm^2-sec, and the initial effective minority carrier lifetime of the base regions, $\tau_{no} = 5 \times 10^{-8}$ sec. These constants will be used in the rest of the investigation.

It will be shown later in this paper that some of the devices studied are useful power switches even though the d-c gain (h_{FE}) of the narrow base transistor has been degraded well below unity. Thus for comparable basewidths the PNPN is inherently more radiation-resistant than the transistor.

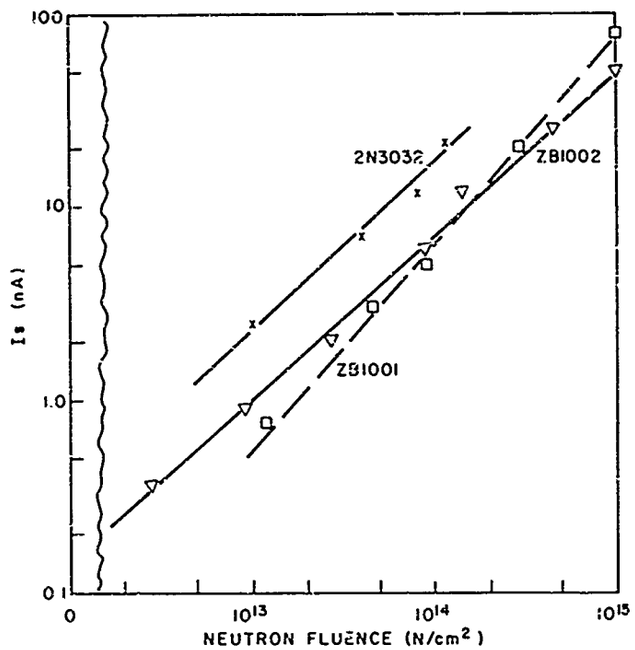


Figure III-4. PNPN Center Junction Saturation Current versus Neutron Fluence

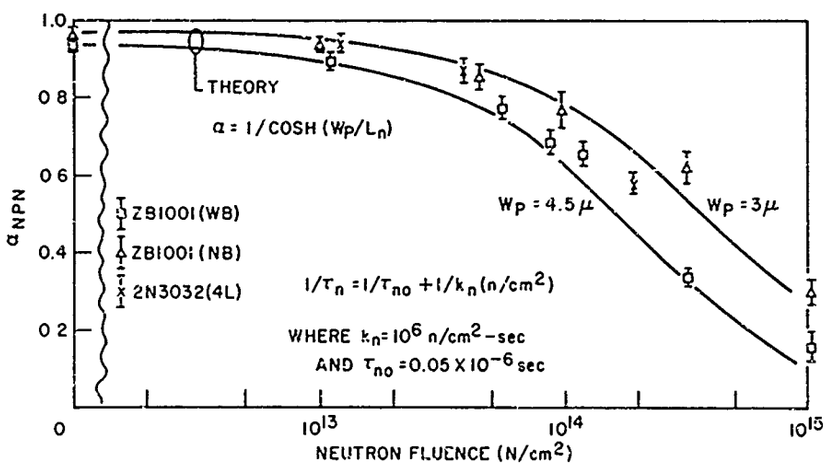


Figure III-5. PNPN Devices α NPN versus Neutron Fluence (I_C = 1.0 ma)

d. Holding Current and Gate Current

Alphas in the vicinity of the turn-on condition having been determined, the variation of holding current and gate current with respect to an increase in the neutron fluence will now be discussed. Figures III-6 and III-7 show plots of the experimental results of the holding currents, which vary approximately linearly with the neutron fluence. This dependence may be attributed to the approximately linear relationship between the alpha, the normalized voltage drop, and the current density in the lightly doped wide base region (N^-), as was shown in Figure III-2. From Figure III-2 it is evident also that devices with larger basewidths require a higher normalized voltage drop V_1 or higher current level to attain the holding current conditions $\alpha_1 + \alpha_2 \approx 1$. This is also reflected in the experimental data of Figures III-6 and III-7. For example, for the same neutron fluence, devices with narrow basewidths such as ZB1001 (15 μ basewidth) have holding currents of about one order of magnitude lower than that of the conventional device such as 2N1765 with a 100 μ basewidth.

Using the experimental value of α_2 from Figure III-5 and $\alpha_1 \cong 1 - \alpha_2$ for the known W_n/L_p ratio, an approximate holding current can be estimated. The result shows an approximately linear dependence of holding current on the neutron fluence, but is found to be an order of magnitude greater than the experimental values. This discrepancy has two possible explanations: first, a localization of the current in the holding condition, and second, change of the minority carrier lifetime with the injection level. The first possibility is supported by the following observation. In the case of ZB1001 ($W_n = 15\ \mu$, $N_d = 10^{15}/cm^{-3}$) at 10^{15} N/cm^2 fluence, the experimental value of the holding current is 12 ma. If the current is uniformly distributed over the emitter area (3.2×10^{-3} cm^2), the corresponding current density is 3.7 amp/cm^2. Assuming that all the increase in "on" voltage (ΔV) at 10^{15} N/cm^2 results from a uniform field in the wide base, then the holding current density must be

$$J = \sigma \frac{\Delta V}{W_n} = .125 \times 1.9/15 \times 10^4$$

$$= 160 \text{ amp/cm}^2 \qquad (5)$$

One may then conclude that if a current density of 160 amp/cm^2 is necessary for holding, then there must exist areas of localized current at this current density, although the average current density is only 3.7 amp/cm^2. Localization of current has also been reported for large W/L ratio PIN devices by Barnett and Milnes.[12]

When the basewidth-to-diffusion-length ratio is large and pure drift model becomes adequate, the approximate holding current can be readily derived from the expression $\alpha_1 \cong \exp(-G/2V_1)$ as

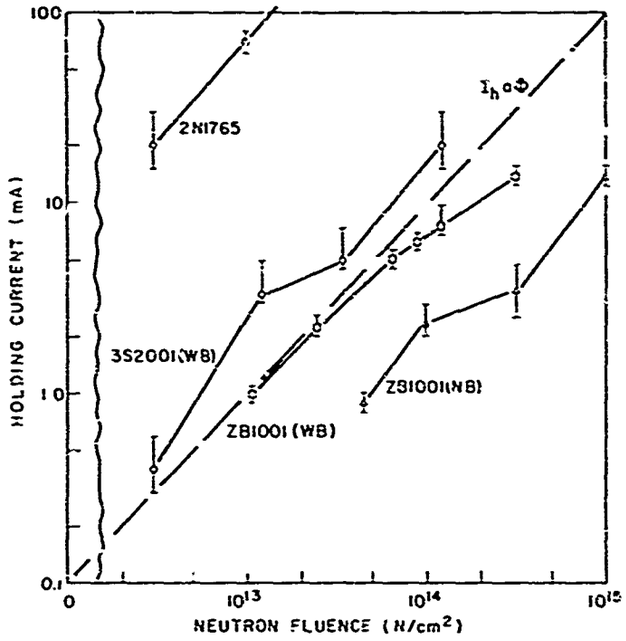


Figure III-6. PNPN Holding Current versus Neutron Fluence (Rgk = ∞)

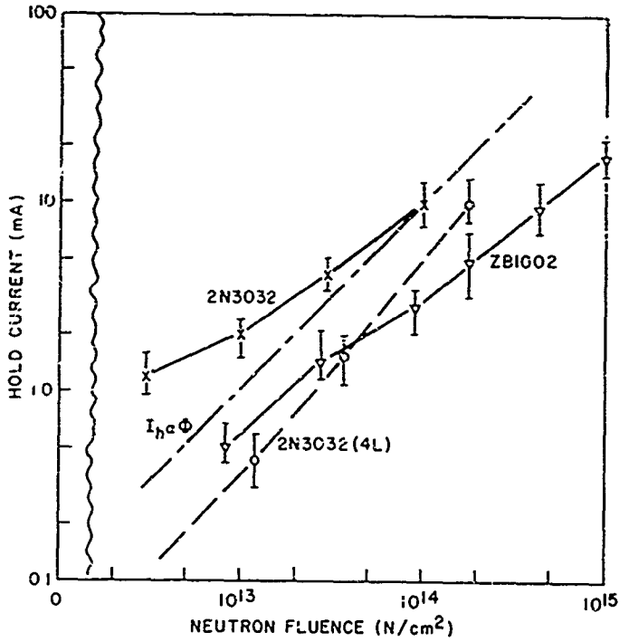


Figure III-7. PNPN Holding Current versus Neutron Fluence (Rgk = ∞)

$$J_h \propto qD_n N_d \frac{W_n}{L_p^2} \frac{1}{\ln\left[1/(1-\alpha_2)\right]} \propto \frac{1}{\tau} \propto \Phi \quad \tau = \text{lifetime} \tag{6}$$

indicating approximately linear relation between J_h and neutron fluence as observed.

Figures III-8 and III-9 show the gate current variation with neutron fluence, where gate current is current into gate used for switching from blocking to "on" states with low voltage (~5v) on the anode. Assuming that in the narrow base NPN transistor the gate current is approximately equal to the turn-on current divided by $h_{FE} = \alpha_2/(1 - \alpha_2)$ and both the turn-on current and the $1/h_{FE}$ vary in proportion to the neutron fluence, one may then deduce that the gate current varies in proportion to the square of the neutron fluence as observed. Thus for devices with large basewidth-to-diffusion-length ratio in the lightly doped wide base and approximately unit ratio in the narrow base region, the gate current can readily be derived as

$$J_g \cong \frac{J_h}{h_{FE}} \cong \frac{qD_n N_d}{2}\left(\frac{W_n}{L_p^2}\right)\left(\frac{W_p}{L_n}\right)^2 \frac{1}{\ln\left\{1+\left[2/\left(W_p/L_n\right)^2\right]\right\}} \sim \frac{1}{\tau^2} \propto \Phi^2 \tag{7}$$

where

$$\frac{W_n}{L_p} >> 1, \frac{W_p}{L_n} \approx 1$$

Note that the quadratic relation comes from the degradation of the minority carrier lifetime in the two base regions. Although both hold and gate currents increase significantly due to radiation bombardment, very large power gains are still obtained with the devices since neither gate nor hold currents are large compared to typical operating currents in the structure.

e. "On" Voltage

Figures III-10 and III-11 show experimental forward "on" voltage versus neutron fluence with the corresponding theoretical value of the wide base voltage drop for the different devices. A good agreement between the calculated values shown in Figure III-3 and the experimental values given in Figure III-10 is obtained for the range of the basewidth-to-diffusion-length ratio W_n/L_p, up to 11. In the ZB1001 with a 15 μ basewidth, the "on" voltage is approximately independent of neutron fluence up to 10^{14} N/cm^2. On the other hand, the conventional triple-diffused 2N1765, with a 100 μ basewidth, is severely degraded at 10^{13} N/cm^2. For much higher W_n/L_p ratios, the exponential relation between the base voltage drop and W_n/L_p ratio predicts a base voltage drop several orders of magnitude higher than the actual

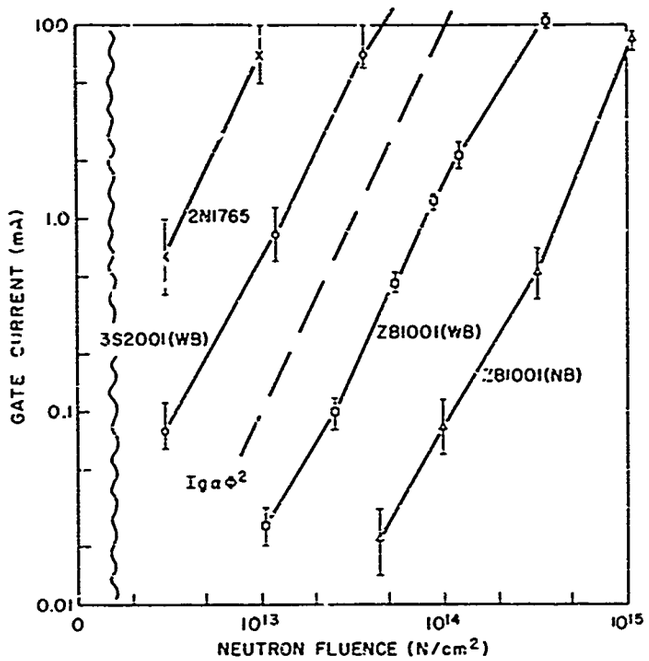


Figure III-8. PNPN Gate Currents versus Neutron Fluence

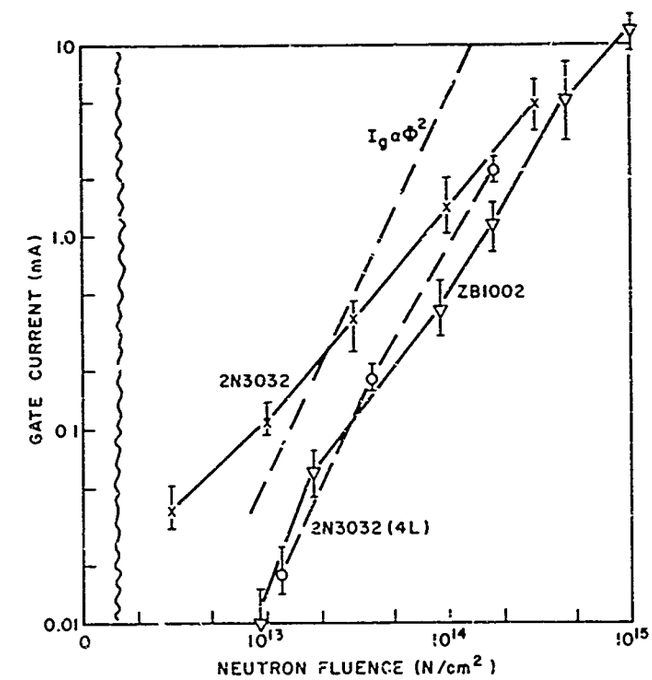


Figure III-9. PNPN Gate Currents versus Neutron Fluence

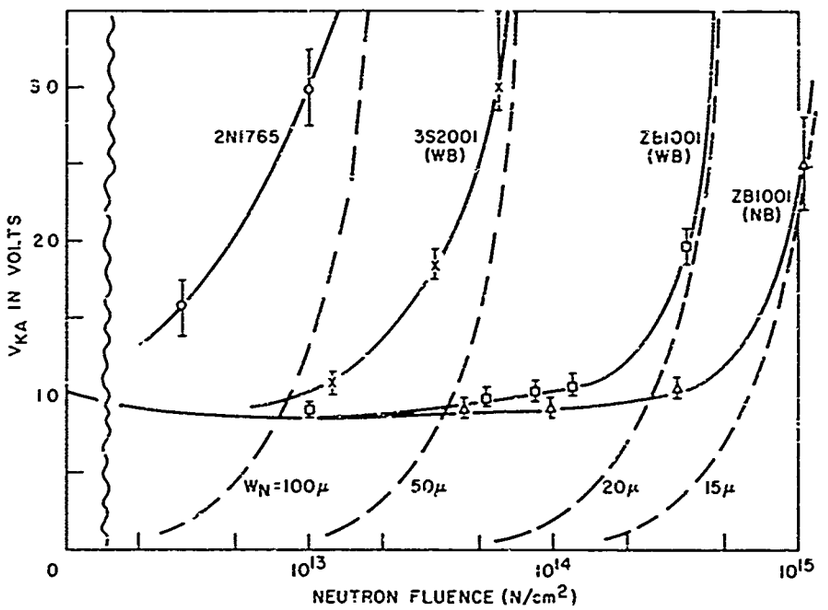


Figure III-10. PNPN Forward "On" Voltage versus Neutron Fluence (500 ma)

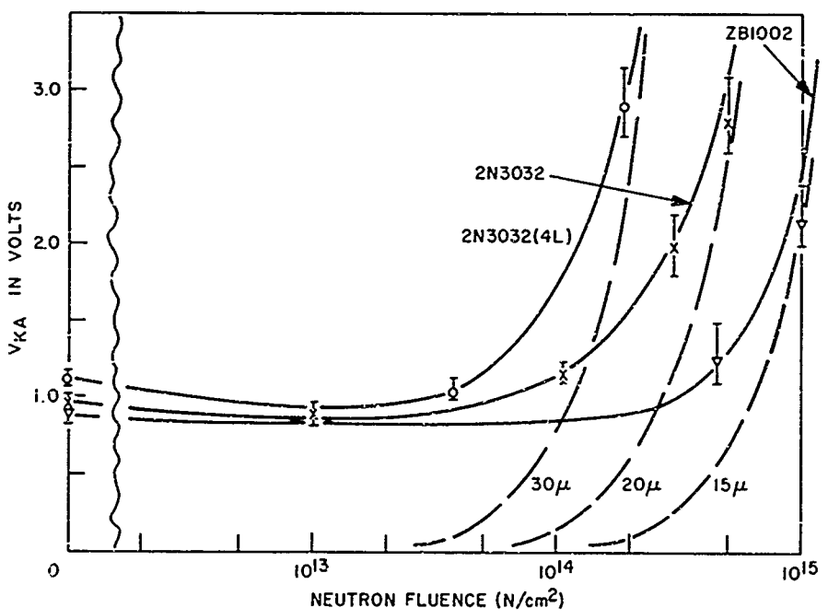


Figure III-11. PNPN Forward Voltage versus Neutron Fluence (100 ma)

experimental values. This suggests a large increase in the effective diffusion length due either to field-aided transport of carriers or to strong injection level dependence of the lifetime.

Figures III-12 and III-13 show experimental comparisons of forward "on" voltage versus current for various devices before and after irradiation. The V-I characteristics show the two expected domains:

1. A voltage independent of current domain. The injected carrier density is greater than the doping concentration, and the conductivity is increasing at the same rate as current density leading to the constant base voltage drop in the wide base region. Since the junction voltage is a logarithmic function of the current density, its effect on the slope of the V-I characteristic is negligible.

2. A higher current domain having nearly a quadratic current voltage characteristic. This may be attributed to the combined effects of the fall-off in the injection efficiencies from the left and the right of the wide base region, and the drift-enhanced mode of the carrier flow at the very high injection level, resulting in a condition where the injected carrier density is approximately proportional to the square root of the current density.

The displacement of the V-I characteristics toward higher voltage with increasing fast neutron fluence reflects the increase in the voltage drop of the lightly doped wide base region as the minority carrier lifetime is degraded.

The tendency for the slope of the V-I characteristics to decrease with the increasing dose is greater for the device with the larger basewidth and suggests that the conductivity modulation is decreasing faster for these devices.

f. Effects of the Neutron Fluence on the Junction Capacitance and the Breakover Voltage

Figure III-14 shows the effect of neutron fluence on the junction capacitance of the reverse biased center junction for a device forward voltage of 5 volts. At neutron fluences greater than 10^{14} N/cm^2, the majority carrier concentration is decreased and the depletion region is widened, causing a decrease in the junction capacitance. The effects are more serious for a device with a higher resistivity base. In the lowest resistivity devices (ZB1002) the changes are relatively small even at 10^{15} N/cm^2.

Figure III-15 shows the effect of the neutron fluence on the forward blocking voltage under two-terminal operation. Preirradiated forward blocking voltages in many of the narrow base devices ranged over wide values, due to the fact that the summation of alphas exceeds unity, and these devices do not possess uniform forward

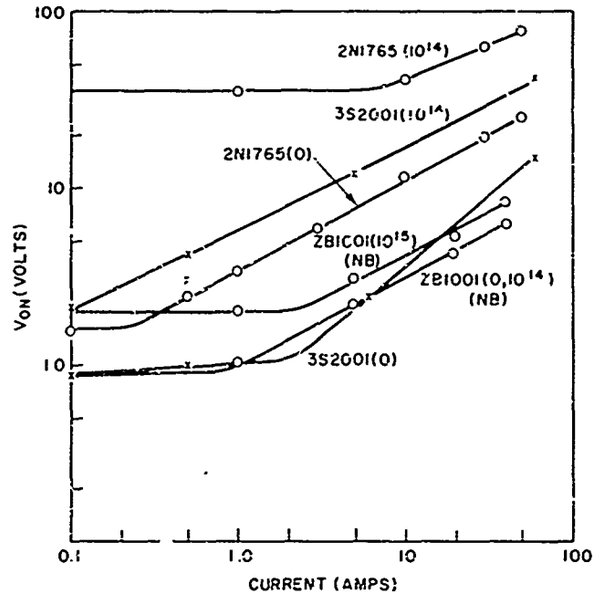


Figure III-12. Forward "On" Voltage versus Anode Current of Irradiated PNPN Devices (0.4 μsec after base pulse, 1 μsec anode pulse)

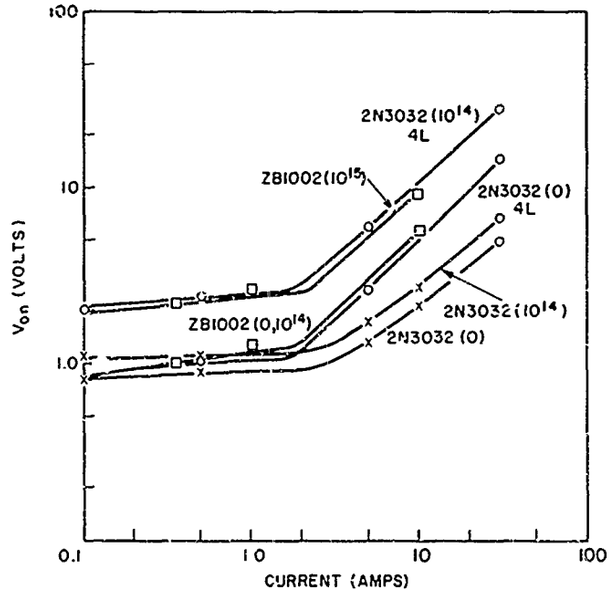


Figure III-13. Forward "On" Voltage versus Anode Current of Irradiated PNPN Devices (0.4 μsec after base pulse, 1 μsec anode pulse)

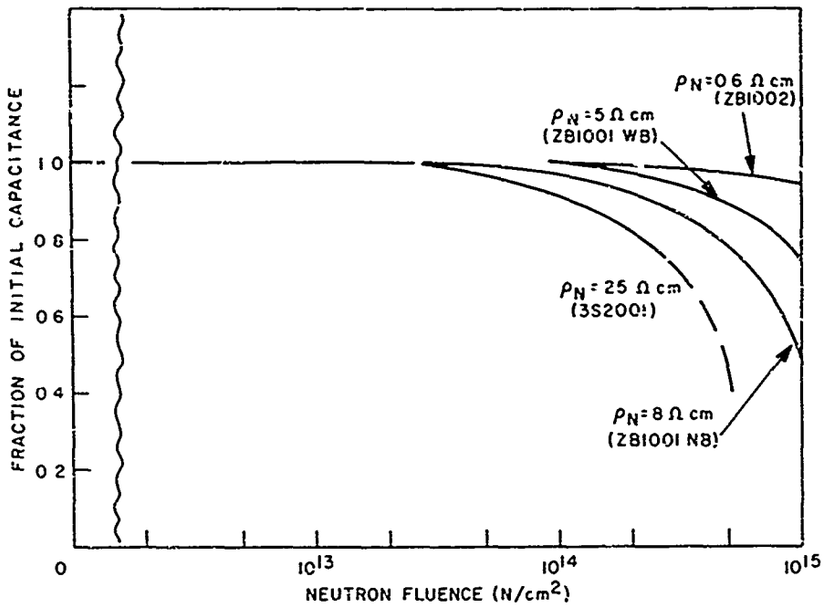


Figure III-14. PNPN Center Junction Capacitance versus Neutron Fluence V_{KA} = 5 volts

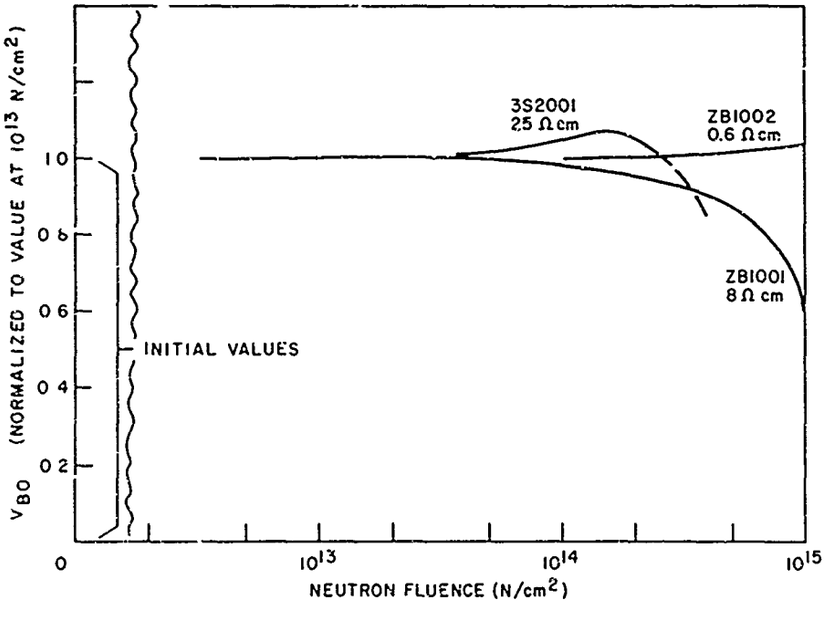


Figure III-15. Forward Breakover Voltage versus Neutron Fluence

blocking characteristics. Therefore forward blocking voltages are normalized to the value corresponding to 10^{13} N/cm^2 fluence where the carrier lifetime has been appreciably degraded in both bases of the narrow base PNPN devices and the reduced alphas introduce a reproducible forward blocking state. For the heavily doped ZB1002 it is observed that the forward blocking voltage is relatively insensitive to neutron fluence but increases slightly as the majority carrier concentration starts to decrease near 10^{15} N/cm^2. For the narrow base ZB1001 with relatively high resistivity, the breakover voltage is determined by the decrease in the effective basewidth as the majority carrier concentration is decreased at higher neutron fluences. Punch-through of the lightly doped N^- base region may occur at very high neutron fluence levels. In the case of the 3S2001 with a 25 ohm-cm base resistivity and a 50 μ basewidth, there is an initial increase for the breakover voltage due to the decrease of carrier concentration; then, as the carrier concentration further decreases, the space charge region widens, punch-through sets in, and the breakover voltage decreases again.

g. Gate Characteristic

Figure III-16 shows an increase in the gate-to-cathode voltage with increasing neutron fluence. This is caused by the decreased conductivity modulation in the gate region and an effective increase in the gate-to-cathode resistance during the "gate

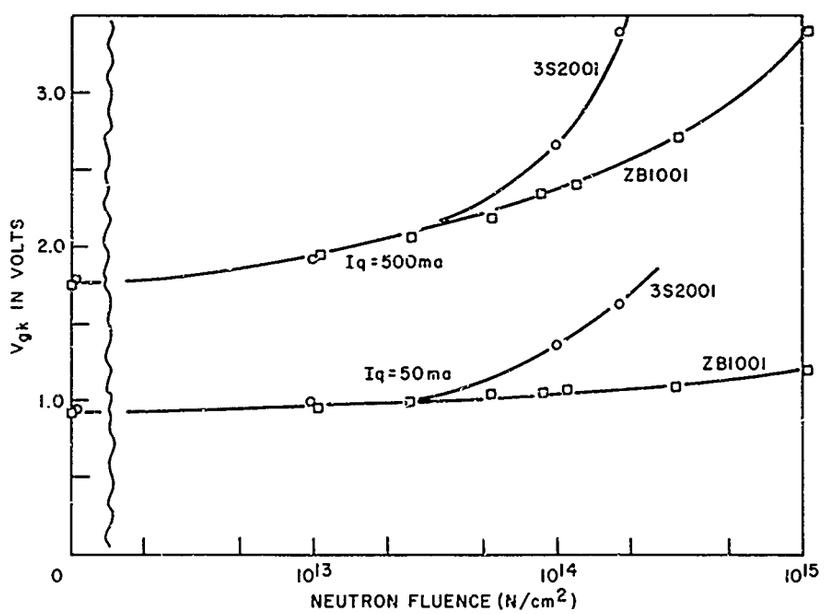


Figure III-16. Gate-to-Cathode Voltage versus Neutron Fluence

turn-on" period. The device with a better interdigitated gate structure, such as ZB1001, is less susceptible to the neutron fluence, since the interdigitated gate structure provides a larger periphery of the cathode-emitter region and lower effective gate-to-cathode resistance. The interdigitated structure also speeds up the lateral propagation of the turn-on state along the emitter junction.

6. DESIGN CONSIDERATIONS

It is now evident that the most severely degraded parameter, and a prime factor in determining the losses and the current handling capability of a PNPN device, is the forward "on" voltage. For optimum radiation hardness and lowest on-voltage, it is desirable that both basewidths of the PNPN be made as small as possible. Since these widths also determine the forward and reverse blocking voltages, a compromise between the forward blocking voltage and "on" voltage must be made. The maximum forward blocking voltage that a PNPN, in two-terminal operations, can achieve is the lowest of either the punch-through voltage or the avalanche breakdown voltage of the reverse biased center junction. The punch-through voltage depends on the particular impurity distribution. In the linear graded junction, the space charge is divided equally between the two sides of the junction. Gaussian types of distribution have the space charge layer mostly in the lightly doped region. However, for a crude estimate of the relationship between breakover voltage, space charge width, and the impurity concentration, a P^+N^- step junction will be assumed for the reverse biased center junction. The relationships are given in Figure III-17. This figure shows, for example, that a PNPN device with a 10 μ basewidth and a doping concentration of 10^{15} cm^{-3} has a punch-through voltage of approximately 80 volts. If the concentration is increased to 3×10^{15} cm^{-3} the junction will avalanche breakdown at 160 volts.

To insure avalanche breakdown rather than punch-through, a thin N^+ layer of high concentration may be inserted between the P^+ region and the N^- base to stop further expansion of the depletion region. In the case of the epitaxial structure, this high concentration layer is deposited first and then the N^- region over it. Transport through the high conductivity region is diffusion-controlled; hence, it must be quite thin. With such a stopping region it is possible to reduce the doping in the N^- region to near intrinsic levels. The electric field in the N^- region is then uniform throughout the region rather than peaked at the PN^- junction, and the breakdown voltage increases to the upper line (marked V'_B) as shown in Figure III-17. The resulting structure, $N^+PIN^+P^+$, has many features in common with the PIN dicde.

Since this structure offers the optimum compromise of blocking voltage and "on" voltage, it will be used in subsequent design considerations. If the $N^+PN^-P^+$

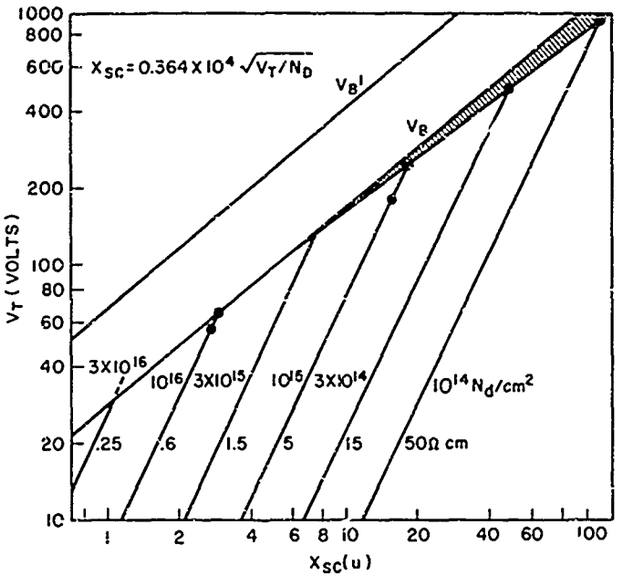


Figure III-17. Space Charge Width versus Reverse Bias for P^+N Step Junction

is used, the blocking voltage capability is reduced by roughly a factor of two provided that the doping in the N^- region is adjusted to yield avalanche rather than punch-through breakdown.

It should be evident from Figure III-17 that high blocking voltage capability is incompatible with radiation hardness. However, the low forward blocking voltage of the radiation-hardened device may be tolerated by operating several PNPN devices in series connection. Assuming a given forward blocking voltage, the conditions for a minimum power dissipation can be calculated from Equations 2 and 3. It can be shown that the optimum condition occurs when the W_n/L_p ratio is equal to 4 (see Appendix A).

This optimum condition occurs when the voltage drop across the wide base is roughly equal to that across the P^+N^- junction, i.e., is of the order of 1 volt. It is possible using this optimum W_n/L_p ratio to relate the blocking voltage of a N^+PINP^+ and the minority carrier lifetime (after irradiation). This relationship is given in Figure III-18. This figure can be used as a design guide for the development of optimally hardened PNPN devices. For example, if it is desired to operate a device in an environment which reduces the lifetime ultimately to 10^{-9} sec (corresponding to a neutron fluence of approximately 10^{15} N/cm^2), the optimum basewidth is 5 μ and the blocking voltage capability of a single N^+PINP^+ device is $\approx$ 160 volts (or

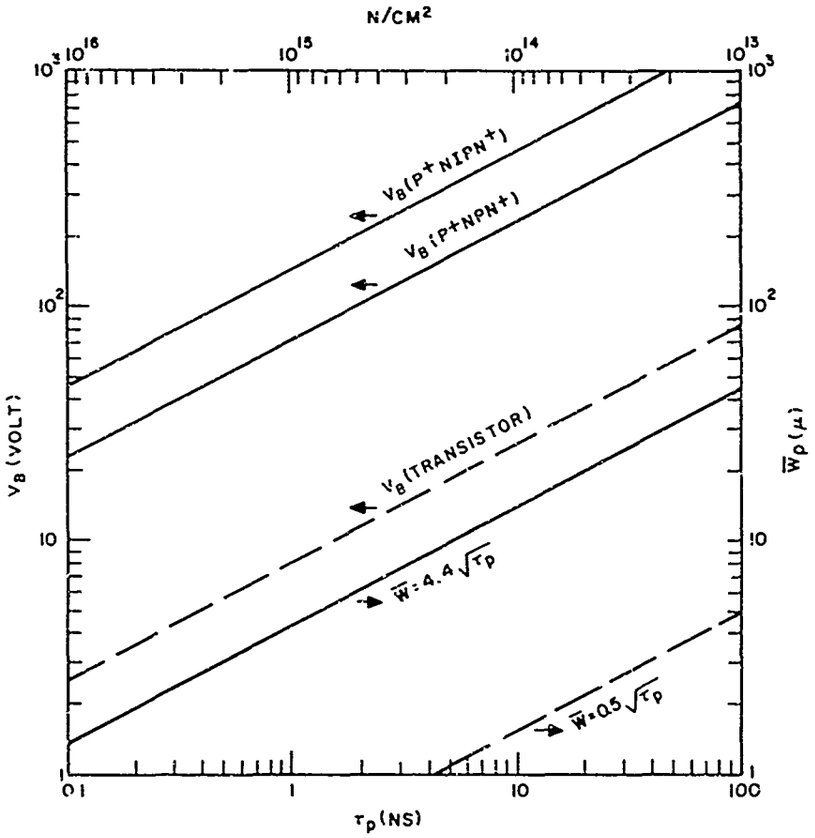


Figure III-18. Reverse Bias and Base Width versus Lifetime for N^+P Step Junction

80 volts for a $N^+PN^-P^+$ structure). In practical devices where structural variations are usually present, blocking voltages roughly one half of these values could be expected. Any required blocking voltage may be achieved then by stacking an appropriate number of devices. Figure III-18 can also be applied to the design of PIN diodes for use in a radiation environment.

From the results given in Figure III-18, a PNPN structure has been evolved for use in neutron environments up to 10^{16} N/cm^2. The structure (Figure III-19) is $P^+NP^-PN^+$ with a P^- thickness of 2 μ and an estimated blocking voltage capability between 30 and 60 volts. The cathode and gate diffusion are interdigitated to minimize effective gate resistance. N^+ regions are diffused into the gate to further minimize this resistance. Impurity concentrations of all regions except the P^- region are sufficiently large that carrier removal will not degrade the performance. Carrier removal in the P^- region will make this region intrinsic but, as previously

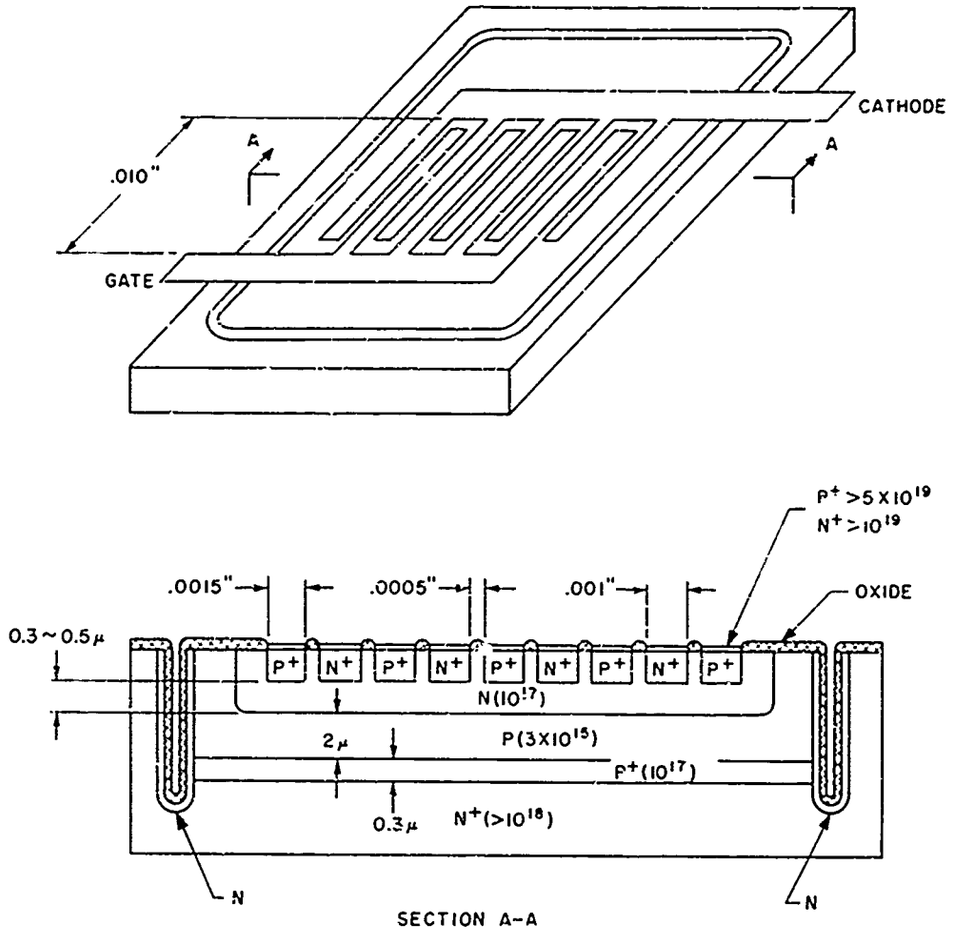


Figure III-19. Junction Structure of an Interdigitated PNPN

pointed out, this is a desirable result. The thicknesses of both N and P regions are such that the degradation of diffusion transport in both does not reduce the sum of the transport factors below unity. Mobility variations from neutron irradiation have been neglected in the design considerations.

This structure should be capable of switching currents of the order of 5 amps with turn-on and turn-off times less than 10^{-9} sec. Because of the narrow bases, the blocking state of this device will not be observed unless the initial lifetimes are drastically reduced either by pre-irradiation or by gold doping.

7. RELATIVE VOLT-AMPERE SWITCHING CAPABILITIES OF PNPN DEVICES AND TRANSISTORS

In designing radiation resistant transistors, as in the design of PNPN's, a compromise between collector breakdown voltage and ultimate device lifetime must be made.[13] This compromise is necessary for the following reasons. Transistors which must function despite short bulk lifetime require narrow basewidths. Figure III-18 shows the transistor basewidth required to yield a current gain of 10 as a function of bulk lifetime. It should be noted that the transistor basewidth required is considerably narrower than either basewidth required in the PNPN. In such narrow base devices, the base resistance is high, emission crowding is severe, and very large current densities enter the collector region. The mobile space charge associated with the collector current will cause the collector boundary to retreat when mobile charge concentration becomes comparable to the majority carrier concentration in the collector. This is the "Kirk effect."[14] To prevent loss in current gain due to the widening of the base, the collector must be heavily doped and its avalanche voltage will be correspondingly low. Figure III-18 shows typical transistor breakdown voltage as a function of lifetime required for $h_{FE} = 10$, obtained on a variety of high-frequency silicon transistors. Again it should be noted that the voltage switching capability of the transistor is roughly an order of magnitude less than that of the PNPN in the same environment.

The PNPN is also superior to the transistor in current switching capability. Since the current in a PNPN is spread uniformly over its cathode area, the current switching capability increases in proportion with the cathode area. In contrast, the current switching capability of a transistor is controlled by the length of the emitter perimeter because of "emission crowding." Thus complex interdigitated emitter-base structures are necessary for high current transistors. For very narrow basewidth transistors there are present-day technological problems, such as material uniformity, which limit the maximum perimeter to the order of 1 cm for basewidths larger than 1 μ and somewhat smaller perimeters for narrower basewidth devices.

As has been pointed out, it is also desirable to interdigitate the cathode-gate structure of the PNPN. Such interdigitation reduces the cathode gate resistance and speeds the turn-on process. To some degree the maximum volt-ampere switching capability (product of maximum current and maximum voltage switching capabilities) is determined by the speed of switching between high voltage - low current and low voltage - high current states. In the case of the PNPN with gate turn-on, the cathode region adjacent to the gate switches quickly, while the remote regions may switch relatively slowly. It is therefore desirable to minimize the lateral distance which the turn-on process must traverse. Once the turn-on process is completed, however, the current is spread over the full area of the cathode. The requirement for rapid

spreading of the turn-on process combined with the maximum perimeter limitation given above will result in a maximum useful area of narrow base PNPN devices of the order of 1 mm^2.

It should be apparent from the foregoing discussion that the volt-ampere switching capability of the PNPN will be significantly greater than that of a transistor designed for the same nuclear environment. Table III-2 compares two high-frequency silicon NPN transistors and one of the PNPN's studied above. All three devices are representative of the state of the art. The narrow basewidth PNPN is capable of switching more than an order of magnitude higher volt-ampere product than narrow base silicon transistors which survive the same environment. The end of life environment for the transistor is taken as the fluence at which its d-c gain h_{FE} falls to 10, and for the PNPN it is the fluence at which the on-voltage, V_{on}, rises to 3 volts. Both of these criteria correspond roughly to the point where the power switching efficiency drops to 90%.

Table III-2

POWER SWITCHING CAPABILITY FOR TYPICAL DEVICES

Device	Basewidth (μ)	I_{max} (A)	V_{max} BV_{CEO} Blocking Volt	P_{max} (watts)	Φ_{max} (N/cm^2) (h_{FE} = 10 or V_{on} = 3 volts)
NPN (SF2585)	1	3	50	150	10^{14}
NPN (RT-6)[15]	0.3	0.1	15	1.5	10^{15}
PNPN (ZB1001)	3 (P) 15 (N)	10	100	10^3	10^{15}

The superiority of the PNPN with respect to the transistor as a power switch at very low switching speeds (i.e., for wide base devices) has been generally recognized.[15] The superiority of narrow base PNPN's to transistors has not been so generally appreciated. Since these narrow base PNPN's are inherently much higher speed devices, it follows that for comparable switching speeds the volt-ampere capability of the PNPN also exceeds that of the transistor. In a recent article E. O. Johnson[16] has shown that PNPN devices approach much closer to the theoretical limiting value of $(V_m I_m X_c)^{1/2} f_t$ than do transistors. X_c and f_t are the device impedance and cutoff frequency respectively. This fundamental superiority is due not only to the absence of emission crowding effects but also to the fact that collector space-charge effects are canceled in the PNPN because holes and electrons enter from opposite sides of the wide base.

The high switching speed capability of the ZB1001 has been employed to make a pulse generator capable of producing pulses 10^{-8} sec in length with an amplitude of 10 amps. The very narrow base design described in the previous section should be capable of producing pulses 10^{-9} sec with an amplitude of 5 amps.

8. SUMMARY

Narrow base PNPN devices exhibit a high degree of radiation resistance. The device parameter which determines the environmental limitations of PNPN's is the forward "on" voltage, which is a sensitive exponential function of the W_n/L_p of the wide base. Consequently narrow base devices have a considerably reduced vulnerability. A reasonably good fit of the observed "on" voltage for a variety of silicon PNPN's exposed to fast neutrons was obtained assuming a silicon lifetime damage constant of 10^6 sec-cm^2/N using an analysis based on the work of Kuzmin.[7]

Design considerations based on these results suggest a device using a P^+NIPN^+ structure which should be capable of surviving nearly 10^{16} N/cm^2 and of switching a volt-ampere product greater than 100 watts. This same structure should also be useful as a high frequency - high power switch. Since such a device functions despite short bulk lifetimes it should also be insensitive to undesired turn-on by a transient ionization pulse. It is estimated that an ionizing dose rate greater than 10^9 rads/sec would be required to turn on this device if the initial lifetime were degraded to 3×10^{-10} sec.

In this study it has been found that PNPN's are markedly superior to transistors as power switches to be used in a radiation environment. State of the art PNPN's have a volt-ampere switching product more than an order of magnitude greater than state of the art silicon NPN transistors. Fundamentally this comes about because the recombination current in the base of a PNPN produces a voltage drop which aids the flow of minority carriers, whereas the comparable recombination current in the bases of the transistor produces a lateral voltage drop, emission crowding, and restricted current- and voltage-handling capabilities.

APPENDIX A

DESIGN CONSIDERATIONS FOR SERIES-CONNECTED PNPN DEVICES

This appendix discusses methods for obtaining, in a series string of PNPN devices, the optimum combination of (a) the number of devices and (b) their basewidths to minimize the total power dissipation in the "on" condition. The initial constraints are the minority carrier lifetime and the total breakdown voltage for the series string.

If the minority carrier lifetime is short enough that the voltage drop across the base of one device is comparable to the algebraic sum of its junction voltages, the power dissipation of the N series-connected devices in the "on" condition is as follows:

$$P_{dis} = NJ \frac{kT}{q} \left\{ \ln \left[\frac{J}{2q D_p \left(P_n / \tau_p \right)} \right] + \bar{\delta}\, e^{\left(W_n / 2\tau_p \right)} \right\}$$

The number N and the basewidth W_n for minimum power dissipation are determined in part not only by the required breakover voltage but also by the type of the forward breakover mechanism.

In P^+NPN^+ devices, one may adjust the doping concentration of the N region for a given basewidth to obtain the avalanche breakdown condition. In order to insure avalanche breakdown rather than punch-through, a thin P^+ layer of high concentration may be inserted between the P^+ region and the N^- base. The resulting structure, $P^+N^+N^-PN^+$, has breakdown characteristics similar to those of a reverse-biased PIN diode, and the breakdown voltage is almost proportional to the basewidth. The blocking voltage may be expressed as

$$V_{block} = NCW$$

where C is a weak function of ionization rates and basewidth.[17]

By the method of Lagrange multipliers, $f(N, W, \lambda) = P_{dis} + \lambda NCW$, the condition for the minimum power dissipation for a given forward blocking voltage can be found by simultaneously solving the two equations, $\partial f/\partial N = 0$ and $\partial f/\partial W = 0$. The condition is

$$\delta\, e^{\left(W_n/2\lambda_p\right)\left(\frac{W_n}{2\lambda_p} - 1\right)} = \ln \bar{J}$$

where

$$\lambda_p = L_p \sqrt{\frac{2b}{b+1}} \quad \text{and} \quad b = \frac{\mu_n}{\mu_p}$$

For the current range $\bar{J} \sim 100$, the width-to-diffusion-length ratio becomes approximately $W_n/L_p \simeq 4$.

Because of the strong dependence of the ionization rate on the electric field, the critical field for a given semiconductor varies very slowly with background doping or impurity gradient.[18,19] Thus, as a first approximation, a fixed critical field intensity of 3.3×10^5 volt/cm was chosen. Then, the breakdown voltage is related to the diffusion length through $W_n/L_p = 4$ as $V_b = E_m W_n = 4\, E_m L_p$ for N series-connected P^+NIPN^+ devices.

A similar argument may be applied to N series-connected P^+NPN^+ devices with avalanche breakdown as the breakover mechanism by adjusting the doping of the base region. The breakdown voltage is then related to the diffusion length through $W_n/L_p = 4$ as $E_m(W_n/2) = 2\, E_m L_p$.

The above results are plotted in Figure III-18; data for a single transistor with a one-diffusion-length base are included for comparison.

REFERENCES

1. F. A. Little and J. Blair, "Study of the Effects of Fast Neutrons on Silicon Controlled Rectifiers," IEEE Trans. on Nuclear Science, Vol. NS-12, No. 6 (December 1965) 64.
2. Gail D. Smith, "Performance of Silicon Controlled Rectifiers in Radiation Environment," IEEE Trans. on Nuclear Science, Vol. NS-13, No. 6 (December 1966) 341.
3. F. A. Leith, "Study of the Effects of Fast Neutrons on Silicon Controlled Rectifiers," Electron Device Conference, Washington, D.C., October 25-27, 1962.
4. V. I. Stafeev, "The Effect of the Bulk Resistance of the Semiconductor on the Form of the Current-Voltage Characteristic of a Diode," Soviet phys. Tech. phys., Vol. 3, No. 8 (August 1958) 1502.
5. J. W. Mayer, R. Baron, and O. J. Marsh, "Observation of Double Injection in Long Silicon p-i-n Structures," Phys. Rev., Vol. 137, No. 1A (4 January 1965) A286.
6. R. W. Aldrich and N. Holonyak, Jr., "Multiterminal P-N-P-N Switches," Proc. IRE, Vol. 46 (June 1958) 1236.
7. V. A. Kuz'min, "The V/A Characteristics of Semiconductor Devices of the P-N-P-N Type in the 'On' State," Radiotekhnika i Elektronika, Vol. 8, No. 1 (January 1963) 171.
8. F. E. Gentry, F. W. Gutzwiller, N. Holonyak, Jr., and E. E. Von Zastrow, Semiconductor Controlled Rectifiers: Principles and Applications of p-n-p-n Devices (Englewood Cliffs, New Jersey: Prentice-Hall, 1964).
9. A. Herlet and K. Raithel, "Forward Characteristics of Thyristors in the Fired State," Solid-State Electronics, Vol. 9, No. 11/12 (November/December 1966) 1089.
10. W. Fulop, "Three Terminal Measurements of Current Amplification Factors of Controlled Rectifiers," IEEE Trans. on Electron Devices, Vol. ED-10, No. 3 (May 1963) 120.
11. D. E. Crees and C. A. Hogarth, "Current Gain in P-N-P-N Silicon Controlled Rectifiers," J. Electronics and Control (GB), Vol. 14, No. 5 (May 1963) 481.
12. A. M. Barnett and A. G. Milnes, "Filamentary Injection in Semi-Insulating Silicon," J. Appl. Phys., Vol. 37, No. 11 (October 1966) 4215.

13. P. O. Lauritzen and D. J. Fitzgerald, "Design Tradeoffs for a Neutron Radiation-Tolerant Silicon Transistor," IEEE Trans. on Nuclear Science, Vol. NS-11, No. 5 (November 1964) 39.

14. C. T. Kirk, Jr., "A Theory of Transistor Cutoff Frequency (f_T) Falloff at High Current Densities," IRE Trans. on Electron Devices, Vol. ED-9, No. 2 (March 1962) 164.

15. A. Herlet, "The Design of Thyristors for Various Fields of Application," Siemens Review, Vol. 32, No. 9 (September 1965) 294.

16. E. O. Johnson, "Physical Limitations on Frequency and Power Parameters of Transistors," RCA Review, Vol. 26, No. 2 (June 1965) 163.

17. J. L. Moll, Physics of Semiconductors (New York: McGraw-Hill, 1964).

18. S. M. Sze and G. Gibbons, "Avalanche Breakdown Voltages of Abrupt and Linearly Graded P-N Junctions in Ge, Si, GaAs, and GaP," Appl. Phys. Letters, Vol. 8, No. 5 (1 March 1966) 111.

19. A. Herlet and H. Patalong, "The Reverse Conductivity of Alloyed Si Junction Diodes" (in German), Z. Naturforsch, Vol. 10a, No. 7 (1955) 584.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author)
Bell Telephone Laboratories, Incorporated
on behalf of Western Electric Company, Incorporated
222 Broadway, New York, New York 10038

2a REPORT SECURITY CLASSIFICATION
Unclassified

2b GROUP

3 REPORT TITLE
Effects of Radiation on Semiconductor Materials and Devices

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)
Final Scientific Report. Period covered: 1 Oct 64 thru 30 Nov 66. Approved 1 Mar 67

5 AUTHOR(S) (Last name, first name, initial)
Wilson, D.K. Cuthbert, J. D.
Mitchell, J.P. Blair, R. R.

6 REPORT DATE	7a TOTAL NO OF PAGES	7b NO OF REFS
31 December 1966	256	102

8a CONTRACT OR GRANT NO	9a ORIGINATOR'S REPORT NUMBER(S)
AF19(628)-4157	Final Rep.
b PROJECT NO. 4608-01	
c DOD ELEMENT: 62405314	9b OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d DOD SUBELEMENT: 674608	AFCRL-67-0068

10 AVAILABILITY/LIMITATION NOTICES

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED.

11 SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY
	Hq. AFCRL, OAR (CRW) United States Air Force L.G. Hanscom Field, Bedford, Mass. 01730

13 ABSTRACT

Results of investigations on the effects of nuclear radiation on semiconductor materials, device surfaces, and devices are discussed.

Radiation damage in gallium phosphide was studied using electro- and cathodoluminescence. Studies were also made of radiative and non-radiative recombination mechanisms in various compound and elemental semiconductors. A non-radiative Auger-type mechanism observed at neutral defect centers appears to explain non-radiative lifetime degradation from both chemical and radiation damage defects.

The Fermi level dependence of the ESR spectrum associated with the phosphorus-vacancy complex was studied in electron-bombarded, phosphorus-doped LOPEX silicon. These studies confirm one assumption that the Si-G8 (E) center is not seen until the Fermi level falls below E_c -0.48 eV.

Experiments to determine the effects of device bias, temperature, and radiation dose rate on surface damage to MOS FET's showed qualitative agreement with a model of positive space charge buildup at traps in the devices' SiO_2 layer. Preliminary thermoluminescence studies to determine the source of traps are described.

Transient recovery phenomena in silicon devices after bursts of fast neutrons were studied. A comparison with the cluster model for neutron damage implies that transient recovery is an electronic, not an atomic process. High-injection-level effects in irradiated transistors and PNPN devices are analyzed and compared with experimental results. Narrow-base PNPN devices are shown to be markedly superior to transistors as power switches in a radiation environment.

DD FORM 1 JAN 64 1473

UNCLASSIFIED

Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
1. Silicon						
2. Surface effects						
3. Radiation damage						
4. Gallium phosphide						
5. Thermoluminescence						
6. Transient recovery						
7. Electroluminescence						
8. Space charge buildup						
9. Semiconductor devices						
10. Semiconductor materials						
11. Radiative recombination						
12. Electron spin resonance						
13. Phosphorus vacancy complex						
14. Non-radiative recombination						
15. Silicon dioxide passivation						
16. Automatic Log I vs V plotting						
17. Cathodoluminescence						
18. Fast-neutron damage						
19. Radiation hardening						
20. High injection level						
21. MOS FET						
22. Transistors						
23. PNPN devices						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, &c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.

Unclassified

Security Classification

Unclassified

Security Classification

14. KEY WORDS		LINK A		LINK B		LINK C	
		ROLE	WT	ROLE	WT	ROLE	WT
1. Silicon	17. Cathodoluminescence						
2. Surface effects	18. Fast-neutron damage						
3. Radiation damage	19. Radiation hardening						
4. Gallium phosphide	20. High injection level						
5. Thermoluminescence	21. MOS FET						
6. Transient recovery	22. Transistors						
7. Electroluminescence	23. PNPN devices						
8. Space charge buildup							
9. Semiconductor devices							
10. Semiconductor materials							
11. Radiative recombination							
12. Electron spin resonance							
13. Phosphorus vacancy complex							
14. Non-radiative recombination							
15. Silicon dioxide passivation							
16. Automatic Log I vs V plotting							

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, &c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.

Unclassified

Security Classification